Editor: André W. Gunst	Date of issue: 2007-02-16 Kind of issue: public	Scope: system Doc.id: LOFAR-ASTRON-ADD-006	
	Status: **final** Revision nr: 5.0		

LOFAR Architectural Design Document of the Astronomical Applications

.

Accepted:	
System Engineering Manager	Program Manager
André W. Gunst	Jan Reitsma

Distribution list:

Group:
Review committee E. van den Heuvel (chair, University of Amsterdam) A. Berg (SARA) W. Brouw (University of Groningen) R. Schilizzi (SKA-ISPO/U. Leiden) C.H. Slump (University of Twente) A.B. Smolders (NXP Semiconductors) E. Stolp (Thales Naval Nederland) LOFAR Project Mark Bentum Ger de Bruyn André Gunst Michiel van Haarlem Hanno Holties Auke Latour Ronald Nijboer Ruud Overeem Jan Reitsma John Romein Gijs Schoonderbeek Michael Wise Albert-Jan Boonstra Stefan Wijnholds

Document history:

Rev	Date	Originator	Ch.	Modification
1.0	2002/03/08	Kjeld v.d.Schaaf	All	Creation
2.0	2002/09/26	Kjeld v.d. Schaaf	All	Updated based on internal review, issue for System Requirements Review
3.0	2003/05/30	Marco de Vos	All	Issue for PDR after major overhaul and extensive internal iteration
3.1	2003/08/28	Marco de Vos / Kjeld v.d. Schaaf	6.2	Update of performance tables

4.0 (draft)	2006/10/13	Kjeld v.d. Schaaf	All	Major update to reflect final design. Re-structured subsystems/baseline design overview. Updated top-level system description.
5.0	2006/02/22	André W. Gunst	All	Major update for system CDR Included system choices and the decomposition of the system with interface descriptions Updated the operational modes Included RFI mitigation and calibration approach

Table of contents:

1 Introduction

1.1 Purpose of this document

This document describes and motivates the top-level architecture of the LOw Frequency Array (LOFAR) instrument and is restricted to the astronomical applications only. The user requirements defined in [1] are in this document reflected into a design which was feasible to implement within a relative short time (about 3 years) and within the constraints of the available funding i.e. a limited budget.

1.2 Applicable documents

Figure 1 relates this LOFAR ADD (Architectural Design Document) with the other architectural documents available. The SRS (System Requirements Specification) [1] describes *what* LOFAR should be in order to realize its mission, while the ADD describes, at a system level, *how* LOFAR will be realized to meet these requirements. This maps the requirements onto available technologies, frameworks and algorithms. The breakdown of LOFAR into subsystems is founded in this ADD. The requirements for subsystems are based on both the SRS and this ADD, since they should realize the subsystem functionality within the architecture defined in the system level ADD.

.

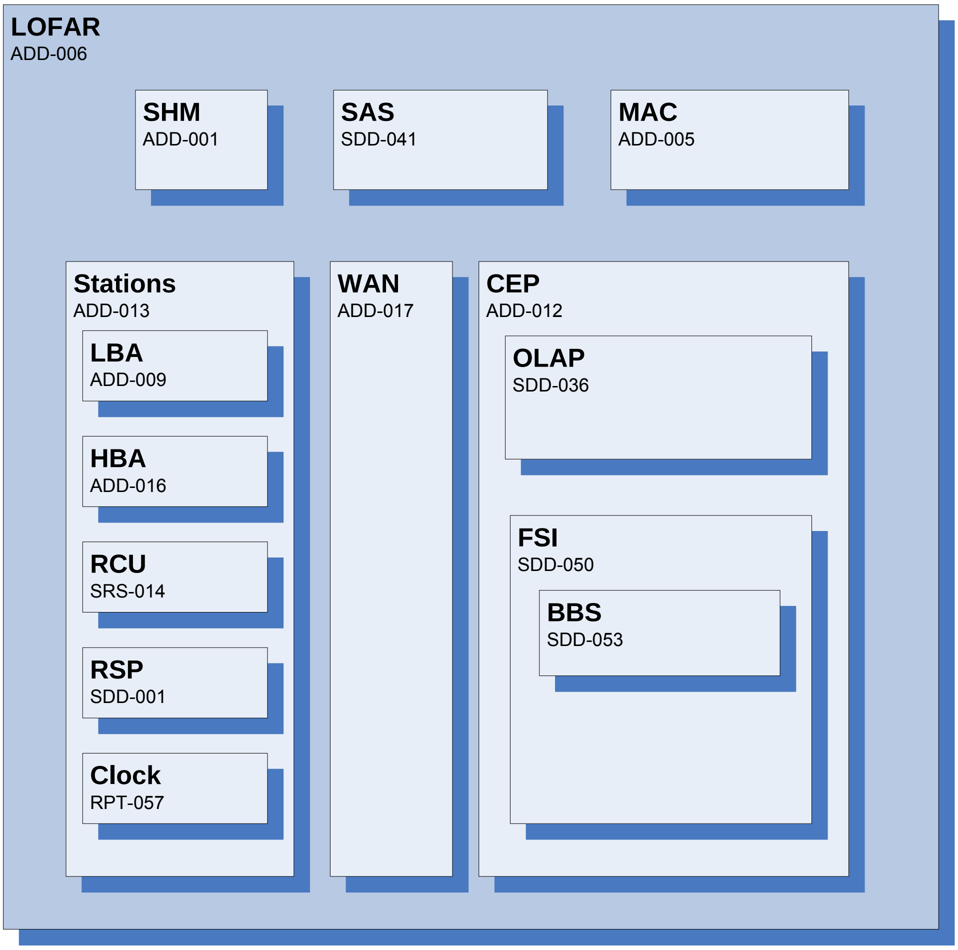

Figure 1 *Dependency relations between this ADD and underlying ADDs available for the LOFAR system.*

1.3 List of Acronyms

ADD Architectural Design Document
BBS BlackBoard SelfCal
BG/L Blue Gene/L series
CCU Central Control Unit
CDR Critical Design Review
CEP CEntral Processor
DOA Direction Of Arrival
DM Dispersion Measures
EPA Embedded Processing Applications
FOV Field Of View
FSI Flagger, Selfcal and Imager
GUI Graphical User Interface
GPS Global Positioning System
GSM Global Sky Model
HBA High Band Antenna
ITS Initial Test Station
KSP Key Science Project
LBA Low Band Antenna
LBH Low Band High antenna
LBL Low Band Low antenna
LCU Local Control Unit
LFFE Low Frequency Front End
LNA Low Noise Amplifier
LOC LOFAR Operations Centre
LOFAR Low Frequency ARray
MAC Monitor And Control
OLAP On-Line Applications and Processing
OTB Observation Tree Browser
PVSS ProzessVisualisierungs- und Steuerungs-System
Rb Rubidium
RCU ReCeiver Unit
RFI Radio Frequency Interference
RCU ReCeiver Unit
RSM Radio Sky Monitor
RSP Remote Station Processing board
SAS Specification, Administration and Scheduling
SCU Station Control Unit
SHM System Health Management
SRS System Requirements Specification
TAB Tied Array Beam
TDS Time Distribution Subrack
TBB Transient Buffer Board
VO Virtual Observatory
WAN Wide Area Network
WSRT Westerbork Synthesis Radio Telescope

2 System architecture

In this chapter the LOFAR system architecture and its decomposition is discussed and motivated. Additionally a functional decomposition is given to partition the functionalities over the subsystems.

2.1 System boundary

The LOFAR system for the astronomical applications is represented in Figure 2[1]. The identified interfaces are

1. Specification: an astronomical user specifies an observation, which configures the LOFAR system in the right mode to observe what is being specified. Typical settings are: beam direction, frequency band, start and stop time, frequency resolution. After the observation is specified and scheduled the LOFAR system starts exactly at the specified time the observation.
2. Sky signals: with this interface the entire sky is meant astronomical signals as well as Radio Frequency Interference (RFI) signals made by human beings. Due to earth rotation and a changing environment, the visible sky changes with time.
3. User data: the LOFAR system generates data products for the astronomer bounded by a specific piece of the sky and bounded in frequency and time. Also metadata is transferred via this interface. The contents of the data are determined through the specification interface.

The astronomical instrument LOFAR is in principle designed and built with four types of applications in mind. These are represented by the four Key Science Projects (KSPs):

1. Epoch Of Reionization [14]
2. Transients [15]
3. Surveys [16]
4. Cosmic rays [17]

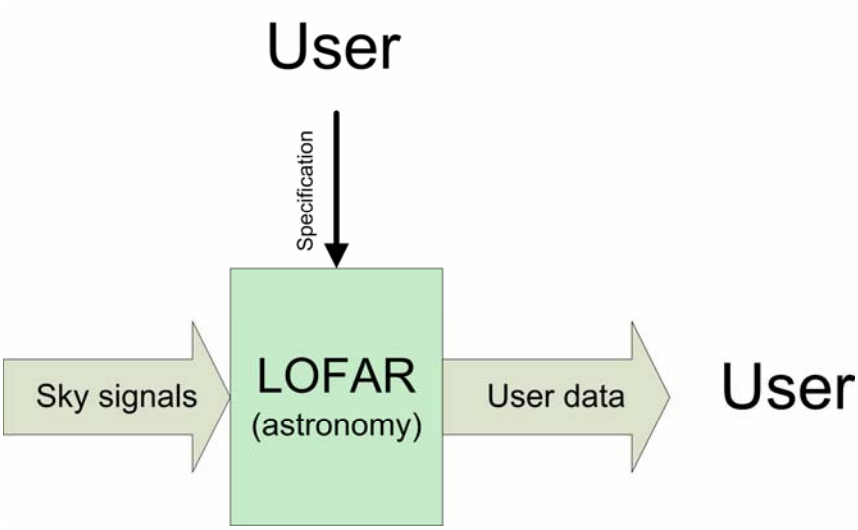

Figure 2 *Interfaces of the LOFAR system*

1 Also a GPS signal is present at the input of the system. Since this follows from an implementation choice it is not included in this figure depicting the primary process. It will appear in the next layer of abstraction.

There is no clear boundary between what functionality is included in the off-line processing of the LOFAR system and what in the off-line processing done by the user. The data volumes involved in LOFAR are so large that many of the data processing operations traditionally done off-line by astronomers must instead be done either in real time or shortly after acquisition, using powerful processing facilities. Essentially, the data needs to be calibrated and processed with sufficient accuracy, so that they can be averaged in time and frequency without significant loss of astronomical information. At that point, data can be stored and exported to an external data centre where the scientists can access and/or further process the data.

In 2006, users already routinely handle datasets of many gigabytes, and during the LOFAR operational phase it will be reasonable to consider exporting datasets containing many hundreds of gigabytes, expanding as time passes. However, visibility data with sufficient spectral and time resolution to image an entire LOFAR station beam at full resolution without time average or bandwidth smearing still implies data volumes of order 10 TByte per observation, so it will be impractical to export experiments involving long integrations with wide fields of view on a routine basis. Therefore, data processing is mainly performed at the LOFAR central processor facility and at the Science Centres, which have high bandwidth connections to the central site and may host additional processing and storage resources. All (user) applications which run on LOFAR hardware are considered to be part of the LOFAR system. This excludes the processing done at the Science Centres or at the users desktop.

2.2 System concept

LOFAR will be the first large radio telescope system, wherein a huge amount of small sensors are used to achieve its sensitivity instead of a small number of big dishes. The main reasons for this are:

- For the low frequencies involved in LOFAR traditional telescopes would be very large and hence costly.
- Pointing can be done electronically, without using moveable parts and hence saving on maintenance costs.
- It enables pointing in multiple directions at the same time.
- It provides operational flexibility (e.g. rapid switching between observations is possible).

Since the concept of LOFAR is so different compared with the traditional radio telescopes, the astronomical science that can be done with it is completely different as well. The penalty of adopting such a new concept is that the way towards its realization is not paved beforehand and introduces hurdles which appear during the engineering and commissioning process. Hence, the instrument itself will remain for a longer period in a research and development phase, to exploit all new functionalities offered by the instrument. That phase will run in parallel with the operational phase.

Despite the new concept the processing steps of a radio telescope remain the same. The general data path for an aperture synthesis array is depicted in Figure 3.

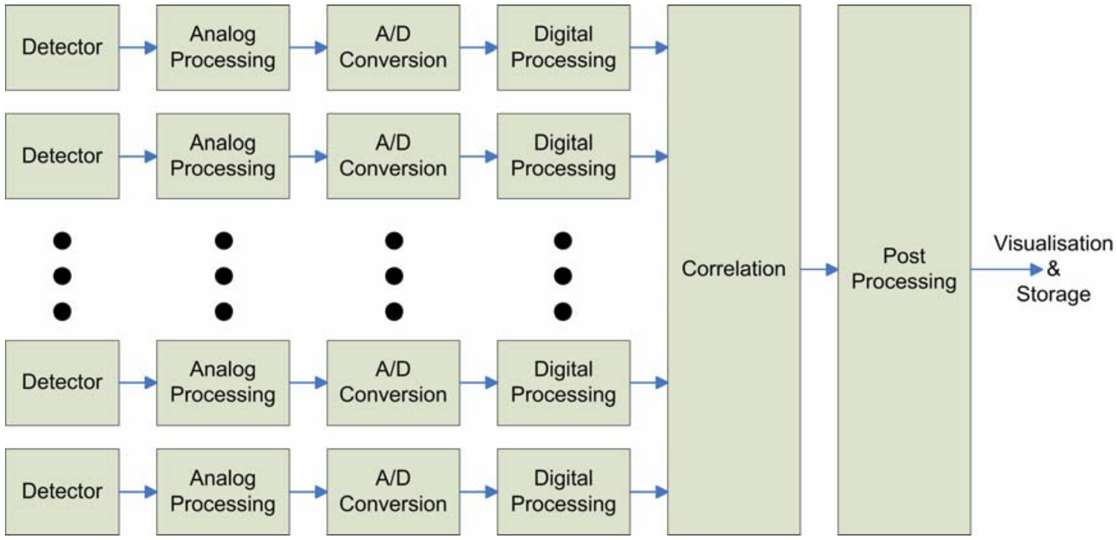

Figure 3 *General datapath of an aperture synthesis array*

In LOFAR, the detector is composed of multiple sensors. Since these sensors convert electromagnetic radiation into electronic signals, the sensors are further referred to as antennas. Ideally, the number of detectors equals the number of antennas to accommodate all sky imaging. However, cost of data transport and processing power limits the number of detectors which can be afforded for a feasible design with a reasonable price. The analog processing shown in Figure 3 covers the (low noise) amplification, filtering, analog signal transport and further signal conditioning functions before the signal is converted into the digital domain (Analog/Digital (A/D) conversion block). From there, the signals are digitally conditioned before entering the correlator. Typical operations in LOFAR digital processing are frequency selection, beam forming, delay tracking and fringe stopping. In the correlator, all signals are correlated with each other to form the cross correlation matrix. Furthermore, the correlation results are calibrated for instrumental and environmental effects. Additionally, known sources are subtracted to enhance the dynamic range. Another post processing task is to transform the correlation products into an image.

For LOFAR, the sensors should be distributed over a large area to achieve an angular resolution of arcsec accuracy with an acceptable UV coverage. All data coming from the sensors should come together in the correlator. So, on the one hand the instrument should be distributed over a large area, while on the other hand all data should come together in a central location. To balance the hardware and operational costs between (1) the equipment in the field, (2) the transport network and (3) the volume of the central systems, multiple antennas (96) are grouped in so called stations. Within such a station the information of all individual antennas is weighted and summed. Such an array of antennas is often called a phased array. By using this technique a spatial selection on the sky is made, which reduces the instantaneous Field Of View (FOV) of each station. The main reason for this is to reduce the total data stream to the correlator.

For at least two Key Science Projects (KSPs) a larger FOV is required in the core of the instrument (2 km in diameter) than at extended ranges. Since there is only a limited distance to bridge from the core to the central processing location, more beams, requiring more bandwidth, can be generated in the core. Hence, in LOFAR two types of stations are distinguished: the remote stations and the core stations.

Finally, the equipment installed in the field and the central systems needs to be controlled. This is taken care of by defining a control layer on top of all systems within LOFAR. The previous discussion is summarized in Figure 4.

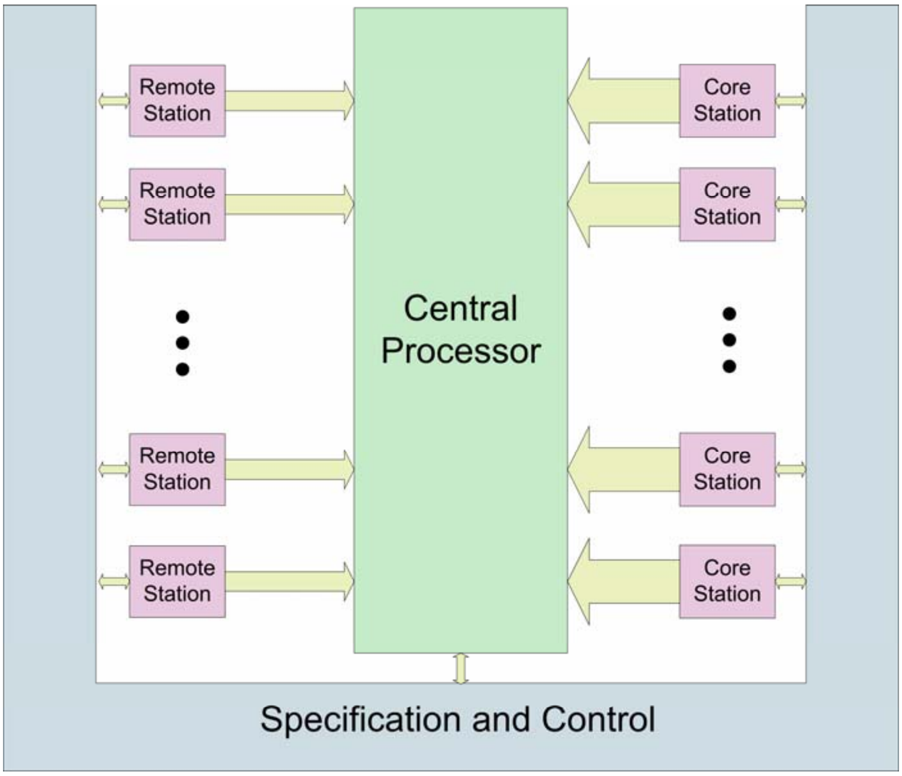

Figure 4 *LOFAR concept composed of remote stations, core stations, central processor, a specification and control layer and the transport between all. The width of the arrows identifies to some extent the amount of data transported.*

2.3 Architectural decomposition of the system

In the LOFAR architecture the following main blocks are identified:

1. Stations [3]. A station selects the sky signals of interest for a particular observation out of the total sky. This process results in one or multiple beams onto the sky. Additionally the station is able to store (raw) antenna data.
2. Wide Area Network (WAN) [7]. The WAN is responsible for the transparent transport of all the beam data (the beam signals on the sky) from the stations to the central processor.
3. CEntral processor (CEP) [4]: The central processor is responsible for the processing and combination of the beam data from all stations in such a way that user data is generated as was specified by the user.
4. Scheduling, Administration and Specification (SAS) [8]. Given the specification, the main responsibility of SAS is to schedule and configure the system in the right mode. Additionally SAS facilitates the possibility to store metadata of the system for a long term and make that information accessible for the user.
5. Monitoring And Control (MAC) [9]. The main responsibility of MAC is to control the system (in real-time) based upon the actual configuration of that moment. Additionally, MAC facilitates the (real-time) monitoring of the present state of the system.
6. System Health Management (SHM) [6]. SHM is identified as an autonomous block to predict and act on failures of the hardware before it actually fails. Ideally it should even pinpoint which system component is the cause of a failure. The reason for considering this block separate from MAC is because of the scale of the system and the percentage of time the system should be effectively operational [1].

The relation between these blocks is depicted in Figure 5. Compared with Figure 2 an extra input signal is drawn. That is the Global Positioning System (GPS) signal which is used to synchronize all stations with each other (in fact this is an implementation decision which is further motivated in Section 2.4.6). In the next sections the interfaces of each block are described.

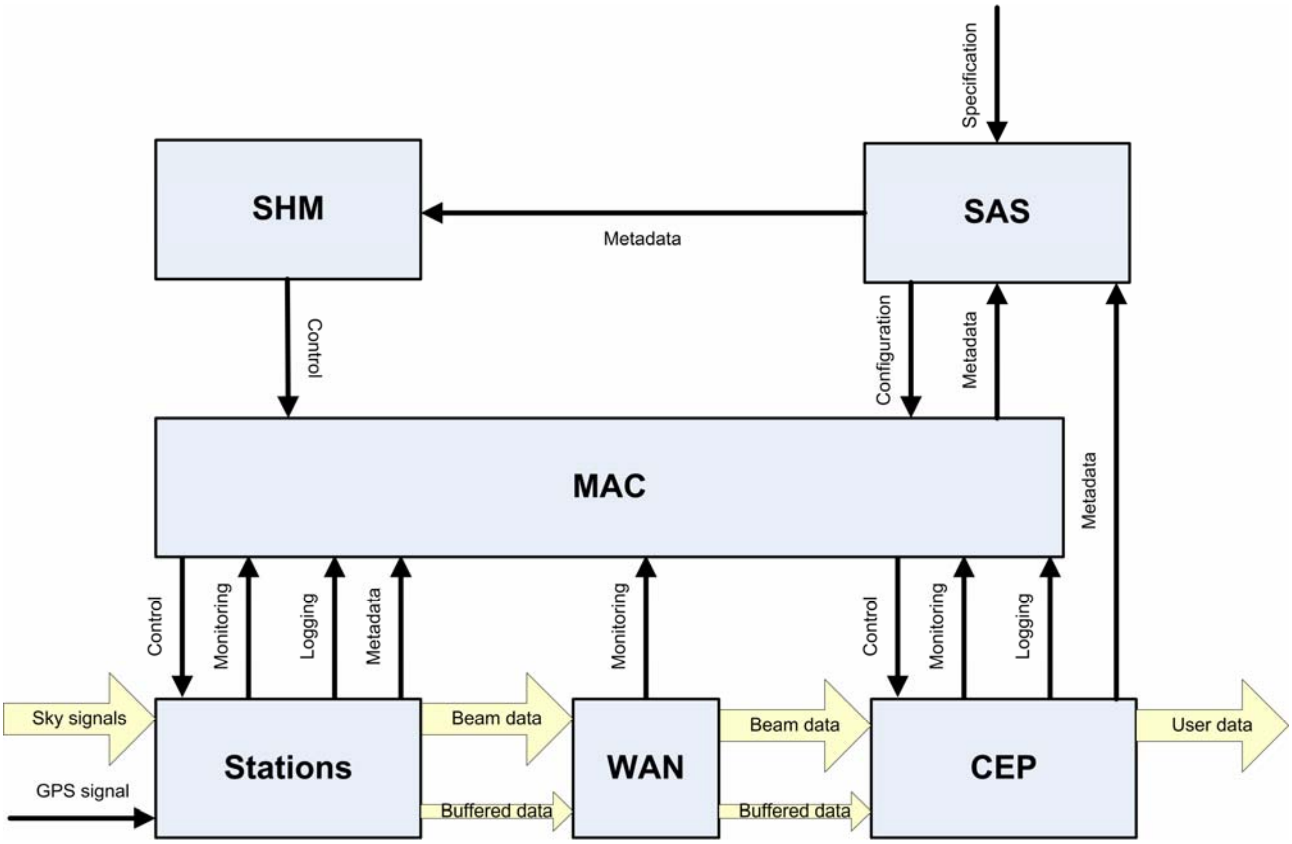

Figure 5 *LOFAR system architecture*

2.3.1 Station interfaces

The station has the following interfaces:

- Sky signals: these are the sky signals received by the station, including man made interference signals.
- GPS signal: this is a signal transmitted by satellites and is used to synchronize the stations
- Beam data: this is a particular selection of the sky signals (or multiple selections) and contains the scientific information of interest. The beam data is a continuous real-time data stream.
- Buffered data: this is either raw or subband data from the antennas, which is stored at the stations to serve specific astronomical applications. Since this data is stored and on request sent to the Central Processor, the data rates involved are much lower than the beam data stream.
- Control: MAC configures the station by this interface to set the station in the right mode.
- Monitoring: via this interface station hardware status information is sent to MAC
- Logging: this information is sent to MAC and contains logging information of the station software. Examples of logging information are warnings or errors that have occurred.
- Metadata: via this channel metadata can be sent to MAC. It contains information about the data itself or information about the environment. Examples are outside temperatures or statistical information of the antenna signals.

2.3.2 WAN interfaces

The WAN subsystem is transparent. The following interfaces can be identified:

- Beam data: this is a particular selection of the sky signals (or multiple selections) and contains the scientific information of interest. The beam data is a continuous real-time data stream.
- Buffered data: this is either raw data coming or filtered data coming from the antennas, which is stored at the stations to serve specific astronomical applications. Since this data is stored and sent to the Central Processor on request, the data rates involved are much lower than for the beam data stream.
- Monitoring: via this line, status information of the WAN hardware is sent to MAC

To be precise, since the MAC subsystem is distributed over the system in fact also the internal MAC communication uses the WAN interface. However, drawing this interface would make the figure confusing.

2.3.3 CEP interfaces

In CEP the following interfaces can be distinguished:

- Beam data: this is a particular selection of the sky signals (or multiple selections) that contains the scientific information of interest. It has to be acquired and processed in real time by CEP.
- User data: astronomical data is the end product of CEP that will leave CEP on request of the user. The definition of the end product depends on the observation mode.
- Control: MAC configures CEP for observations and for example starts / stops observations.
- Monitoring: CEP sends out status information about the CEP hardware and the CEP processes to MAC.
- Logging: CEP sends logging to the logging system which is part of MAC.
- Metadata: During observations, CEP sends metadata to the metadata collection system in MAC (eg. statistics)

2.3.4 MAC interfaces

In MAC the following interfaces are present:

- Control (station): MAC configures the station through this interface to set the station in the right mode.
- Monitoring (station): via this interface station hardware status information is sent to MAC
- Logging (station): this is information is sent to MAC and contains logging information of the station software. Examples of logging information are warnings or errors that have occurred.
- Metadata (station): via this channel metadata can be sent to MAC. It contains information about the data itself or information about the environment. Examples are outside temperatures or statistical information of the antenna signals.
- Monitoring (WAN): via this line status information of the WAN hardware is sent to MAC
- Control (CEP): MAC configures CEP for observations and for example starts / stops observations.
- Monitoring (CEP): CEP sends out status information about the CEP hardware and the CEP processes to MAC.
- Logging (CEP): CEP sends logging to the logging system which is part of MAC.
- Configuration (SAS): SAS translates the specification into a configuration of the system and sends this information to MAC via this interface.
- Metadata (SAS): this interface is used to transport the metadata of the stations to SAS for inspections or for “long” time storage of the metadata (beyond the lifetime of the observation).
- Control (SHM): in case early failures are detected by the SHM system, this interface is used to update MAC with this information in order to take proper action before the components actually fail.

2.3.5 SAS interfaces

In SAS the following interfaces can be identified:

- Specification: through this interface, the user specifies an observation. The level of detail can vary for each application or user.

- Configuration: SAS translates the specification into a configuration of the system and sends this information to the system via this interface.
- Metadata (MAC): this interface is used to transport the metadata of the stations to SAS for inspections or for “long” time storage of the metadata (beyond the lifetime of the observation).
- Metadata (CEP): A direct metadata interface from CEP to SAS is used to store the metadata of CEP for a long term.
- Metadata (SHM): this interface is used to provide SHM with metadata. Primarily metadata concerning hardware is of use for the SHM.

2.3.6 SHM interfaces

In the SHM the following interfaces are distinguished:

- Metadata: this interface is used to get metadata from the system.
- Control: in case early failures are detected by the SHM system, this interface is used to update MAC with this information in order to take proper action before the components actually fail.

2.4 Functional decomposition of the system

In this section the decomposition of the required functionalities over the main subsystems is explained. One of the general rules that are applied is that the amount of hardware in the field is minimized, since maintenance and operational costs of equipment in the field is much higher than equipment at a central location.

First, the global functional data path is discussed to motivate why certain design choices have been made. Based on those design choices, a more detailed functional data path is given at the end of this section.

2.4.1 Global functional data path

The functional signal flow, given the fact that beamforming is essential to reduce the data (Section 2.2) is depicted in Figure 6. The signal flows from left to right, starting off with the antenna receiving the electromagnetic signals and these being converted in to the electrical domain. One of the main functions of the receiver is to amplify and select the frequency of the signal, preserving the signal to noise ratio required. Also, the conversion of the analog signal into a digital signal is part of the receiver.

In interferometry it is important to keep the signal paths equal in (electrical) characteristics (because in fact differences between signals received are measured). This also applies to the signals before beamforming. Any difference in gain or phase introduced prior to the beamforming operation will degrade the signal to noise ratio (here defining “signal” as the signal of interest, the sky noise, and the “noise” as the noise generated by the system). For these reasons early sampling and digitization is preferred and therefore placed before the beamformer in the design shown in Figure 6. In the beamformer, the first spatial selection is made by applying time differences between the receptors in the station field and subsequently summing those signals.

A second compensation for time differences between the signals from the station fields is applied prior to correlation by phasing up the station signals for a certain observation direction (second step of making spatial selectivity). In the correlator the station signals are combined and correlated with each other, resulting in the full correlation matrix. In the post processing block all “after correlation” operations are assigned to, i.e. calibration and imaging.

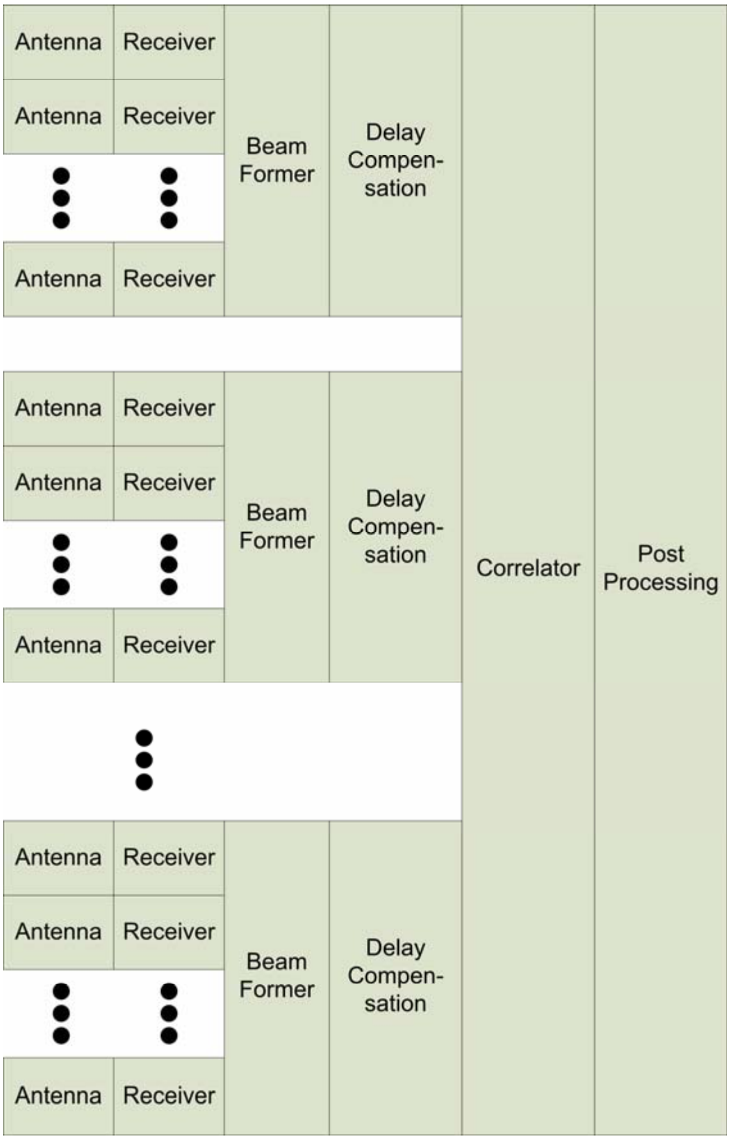

Figure 6 *Global functional LOFAR data path*

In the next sections some of the blocks in Figure 6 are discussed in more detail to motivate the architectural choices leading to the design of the current LOFAR system.

2.4.2 Correlator

It is important to start the functional decomposition with the correlator concept, since the choice for that touches almost all subsystems. In classical radio telescopes an XF correlator was generally used, meaning that first the correlation and integration of the signals was done in time domain (X) where after the Fourier transform (F) was accomplished to get a cross power spectrum out of the correlator. This is still an economically attractive technique for radio telescopes with a limited number of antennas (input signals to the correlator). However, for LOFAR an FX correlator (first Fourier transform and then correlating the resulting channels) is favorable in terms of processing at the expense of data transport (the signals must be regrouped per channel instead of per antenna, resulting in a transpose operation). Using only a Fourier transform in the FX correlator leads to a significant amount of leakage between the channels. Therefore it was chosen to use filter banks before the correlator. This architecture is also known as an HFX (Hybrid FX correlator) architecture.

2.4.3 Beamforming

The antennas in the station form a phased array, producing one or many station beams on the sky. Multi-beaming is a major advantage of the phased array concept. It is not only used to increase observational efficiency, but may be vital for calibration purposes.

To form a phased array at station level, the analog antenna signals must be delayed and added to form a beam on the sky. Moreover the beamformer should be able to track sources on the sky and be flexible in exchanging beams for bandwidth.

The beamformer can be implemented by using true time delays or by applying phase shifts on "narrow" subbands. For the last solution an error is made at the edges of each subband (Figure 7), since the phase is frequency dependent and only one phase can be set per subband.

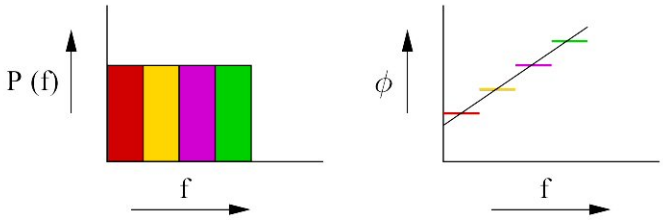

Figure 7 *Illustration of 4 subbands and the error which is introduced by approximating the time delays by phase shifts per subband (the black line on the right hand side is the ideal phase)*

The reason for discussing this topic here is that for the correlator architecture a certain frequency resolution is required (following the specification in [1]) and for the beamformer a certain (other) frequency resolution is required as well (if implemented by phase shifts). This led to the decision to implement the station beamforming with phase shifts after a first stage filterbank which realizes a frequency resolution sufficient for the beamforming operation. A second stage filterbank will make an even higher frequency resolution which is required before the correlator. Another reason to split the filter banks up is that the first stage filter bank is more expensive than the second stage filter bank because the first stage filter bank operates per antenna while the second stage filter bank operates on beams. Since no extra significant data reduction will be done after the second stage filterbank, that functionality will be implemented in the central systems.

2.4.4 Delay compensation

By delay compensation is meant the additional time delay to be applied to the signals from the stations necessary to coherently correlate the station signals. In practice the delay compensation is done in multiple steps, dependent on where in the system a certain time (samples) or frequency resolution is available.

LOFAR should be flexible in the sense that capacity used for beams and for bandwidth can be exchanged [1]. Therefore, it is costly to implement the delay compensation at antenna level because that would force the design from that point onward in the datapath to accommodate hardware for the maximum number of beams and with the widest band required (certainly it is not a trade between beams and frequency anymore). From a delay compensation point of view of course the delay compensation at antenna level would be more favorable (since the time resolution at the A/D converter is high).

ASTRON

To balance between what is necessary and what is affordable, the choice has been made to implement the delay compensation after the beamforming (thus it is implemented only once per beam and not once per antenna). As discussed in the previous sections, two filter banks will be used to make the spectral resolution required for the correlator. The first filter splits the band up in subbands, while the second filter chops the subbands down to channels. The time resolution of the output samples of the first filter can be used to accomplish true time delay compensation. After that, a residual error remains still to be compensated for. This will be done by applying phase shifts to the channels coming out of the second filter bank (also referred to as fringe stopping). Applying only delay compensation after the second filter bank leads to significant decorrelation of the signals. This is calculated and motivated more in detail in [12].

2.4.5 Control

In the data processing pipeline of LOFAR, real time control is required to set the instrument in a certain state at a defined time. Furthermore, the instrument needs to be able to switch fast between modes [1] and be able to track sources within a definite interval. Hence, a distinction is made between real time control during data taking and processing on the one hand, and control prior to this phase (mainly specification) and after that phase (mainly inspection) on the other hand. This separation is motivated by the different types of database technology and software design issues related to real-time operation requirements.

The separation of the various control functions is as follows: the SAS subsystem takes care of the specification and configuration of observations and instrument settings. The MAC subsystem is responsible for the operation of the instrument and the execution of observations, while collecting meta-data about those operations and observations. Finally, the SAS subsystem is used to provide an interface to the users for the collected meta-data and possible snapshots to inspect the observation performance and quality. The process from observation proposal to the user data product is further detailed in Section 6.2.

At system level the choice has been made to control LOFAR centrally so that information is collected (and accessible) in a single place as much as possible. However, one of the requirements is that the stations should be able to function for at least one hour autonomously [1]. Hence, in each station a Local Control Unit (LCU) is present which controls the complete station (the LCU is also referred to as Station Control Unit). An LCU close to a station is also necessary because some hard real-time requirements should be met, e.g. updating the beamformer weights in order to track sources. All LCUs within LOFAR are mastered by so called Central Control Units (CCU). In [8] and [9] the control architecture is discussed in more detail.

2.4.6 Time synchronization

In interferometry, it is important to synchronize all stations with each other since the coherency between the sky signals as received by all the stations is crucial. In classical telescopes the complete system is synchronized by distributing the Local Oscillator signals, clock signals and synchronization signals from one central point. However, LOFAR is too large to adopt such a solution for an acceptable cost. Therefore, it was decided to generate a clock, linked to the GPS signal, for each station. More details about the station synchronization can be found in [44].

All antenna signals within a station and the processing thereof are synchronized. However, because it should be possible to transport the data over a fiber infrastructure not owned by LOFAR, the data is asynchronously sent on the WAN and re-synchronized again at CEP by applying time stamps to the data.

2.5 Detailed functional data path

The choices motivated in the previous sections lead to a detailed LOFAR signal block diagram, as depicted in Figure 8. In the block diagram, a third dimension is introduced to represent the splitting in frequency. This is done to clarify that the operations after the filter bank are independent from subband to subband. Hence, one beamformer is required per subband. This applies also for the correlator. The correlations are independent from channel to channel. The independency gives a certain freedom in the design. Not represented in the diagram is the WAN subsystem since that should act transparently (it is situated on the boundary between beamformer and delay tracking).

Figure 8 *Detailed LOFAR signal data path. The third dimension represents the frequency domain. For example the first filterbank splits up the digital signal into subbands, while the second filterbank splits those subbands further up into channels.*

2.6 Top level system specifications

The following table lists the top level specifications for the array.

Table 1 *Top level specifications*

System parameters		Value
Frequency range		30 – 80 MHz (low band) 120 – 240 MHz (high band)
Polarisations		2
Bandwidth		32 MHz
Spectral channels		42240 52736 in mode III[2]
Stations		32 in core 45 remote
Baseline length		100m to 100 km
Baselines		2926 Full stokes
simultaneous digital beams (full array)	Full array	Configurable between 1 beam of 32 MHz and 8 beams of 4 MHz
	Central Core	24 beams of 32 MHz
Digital signal paths		14784 (2 polarisations x 96 channels per station)
Sample bit depth		12 bit
Correlator capacity		99 10^{12} Correlations/sec
Tied array beamformer capacity		128 beams (full array)
Storage capacity		5 days raw data 1 month reduced data
Correlator data export capacity		20 Gbit/s (200 TByte/day)
Performance parameters (general)		
Spectral resolution		0.76 kHz (0.61 kHz in mode III[2])
Correlator dump time		1 second
Tied array beams		128 beams

[2] The station system can be operated with two different sample clock rates. The frequency range between 190 and 210 MHz is always observed with the lower clock rate, in mode III, resulting in more but smaller frequency channels. See section 3.2 for details.

3 Subsystem Description

This Chapter presents globally the baseline design for LOFAR. The baseline design describes how functions in the dataflow are assigned to subsystems and details the way in which the function of each subsystem is realized.

3.1 Subsystems breakdown

The LOFAR system is decomposed in subsystems, each fulfilling a specific task in the instrument.
The subsystem breakdown is shown in Figure 5. After the introductions in this section, the subsystems are described in more detail in the remaining sections of this chapter.

- **Station systems**
  In the stations the electromagnetic field received by the station antennas is converted into signals with spatial and spectral sensitivity in the digital domain. Two antenna systems are used to be able to observe in two distinct frequency bands. Both antennas include amplifiers to drive analog connections to the signal processing chains. The compound high band array also includes an analog beamformer. The signal processing chain starts with amplification and analog filtering of the antenna signals after which the analog signals are converted to the digital domain. The signal processing chains continue for each receiving element (antenna) with digital filters after which coherent spatially sensitive signals are created in the beamformers. The station output is a time series of voltage beams, although intermediate signals can also be made available for specific observations.
  The signal chains from the remote stations and the core stations use the same signal processing design, but the core stations are equipped with additional hardware to create more beams, thus being able to cover a larger field of view.
- **Wide area network**
  The wide area network transports the station data transparently to the central systems. The control information is sent via this network as well.
- **Central processing**
  At the central site, processing facilities and network infrastructure are available to handle the data streams coming from the stations.
  These facilities are roughly divided in an on-line and an off-line part. The on-line facilities are used during the observation to process in real-time the data streams from the stations. Such operation typically consists of delay compensation, correlation and ionosphere calibration. The resulting data stream from this on-line processing is stored and serves as input data for the off-line processing tasks. These off-line tasks are typically self-calibration, image creation and image analysis / visualisation.
- **Specification and scheduling**
  The instrument control software components fulfil the major task to execute observations on the complete (distributed) instrument. The life cycle of an observation starts in the proposal and specification tools and ends when all data products are archived and exported, from which time onwards the observation data may have a continued life in data archives/database systems, for example at a Science Centre. At this stage, the data may be accessed e.g. through virtual observatory (VO) applications.
- **Monitoring and control**
  After the specification of an observation, the instrument usage is scheduled, resulting in a schedule that is handed over to the control and monitoring package. At the specified time, the measurements will be executed on the specified subsystems of the instrument, in general resulting in a raw dataset being stored at the central storage.
- **System health management**

The system health management subsystem continuous analyses the performance of the instrument and updates the resource availability databases used by the scheduler and monitoring and control packages.

3.2 Station systems

The remote station baseline design consists of 96 Low Band High (LBH) antennas optimized for the 30-80 MHz frequency range and 96 High Band (HB) compound antenna arrays optimized for the 120-240 MHz frequency range. The High Band compound antenna is built up from 16 antenna elements, which are combined through analog beamforming. Both the low band and combined high band antenna signals are pre-filtered and amplified locally prior to transportation over coaxial cables to a central location within a station, as illustrated in Figure 9. Additionally the receiver accommodates for a third antenna input operating in the 10-30 MHz band. This antenna is referred to as the Low Band Low (LBL) antenna and is not developed right now.

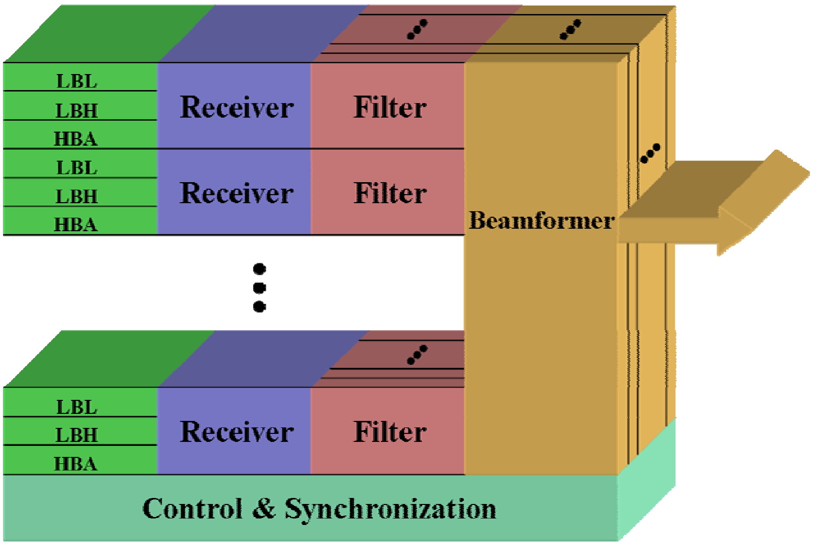

Figure 9 *Station architecture block diagram. The design covers three antenna systems which are connected to the receiver unit where analog filters dedicated for each antenna, can be selected. After digitisation, the antenna signals are filtered. Further beamforming is implemented per subband as indicated by the third dimension.*

At the central station location the receiver unit selects one out of three antennas. After selecting an antenna, the signal is filtered with one of the integrated filters. These filters select one of the four available observing bands. After filtering, the signal is amplified and filtered again to reduce the out of band noise contribution. A pre-amplifier in front of the A/D converter converts the single ended signal into a differential signal prior to A/D conversion. For the receiver a wideband direct digital conversion architecture is adopted. This reduces the number of analog devices used in the signal path. The A/D converter converts the analog signal into a 12 bit digital signal at a maximum sampling rate of 200 MHz. To fill the gaps in between the Nyquist zones, a sample frequency of 160 MHz can be chosen as well. The Nyquist zones I to III of the A/D converter with a sample frequency of 200 MHz and 160 MHz respectively are depicted in Figure 10.

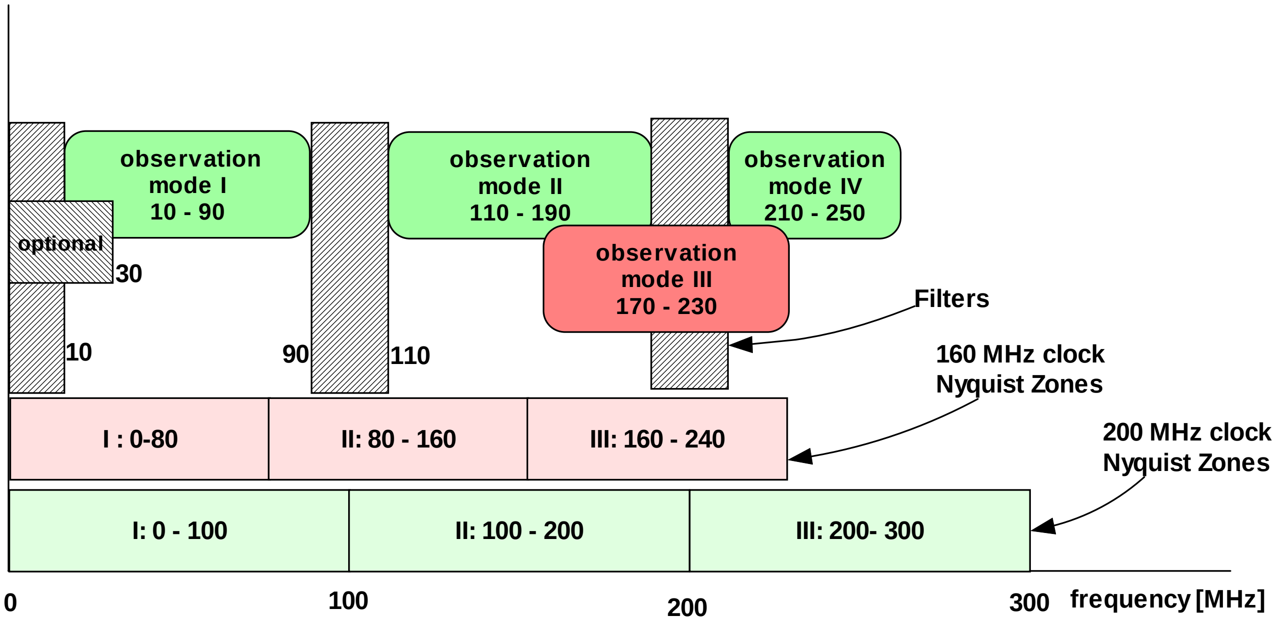

Figure 10 *Selection of Nyquist zones is used to select the observed band in the station.*

After analog to digital conversion, the band is split into 512 equidistant subbands, which are on a fixed grid. The negative part of the original spectrum is omitted, i.e. the real input signal is from here on represented by complex signals.

Each subband signal is decimated with a factor of 1024 after filtering. Hence, the clock rate after filtering is reduced to 195 kHz and 156 kHz respectively for the 200 MHz and 160 MHz input sampling rates. After filtering, specific subbands can be selected. The selection is controlled centrally at the station. The selected subbands have a maximum total effective bandwidth of 32 MHz per polarization. This bandwidth is matched to the current capacity of the central processor.

To form beams, the antenna signals are combined in a complex weighted sum. This is done with independent beamformers for each subband. A beam originating from a specific subband is defined as *beamlet*. The weights necessary for the beamformer are calculated in the local control unit and are sent to the beamformers. In addition, statistical measurements are performed at the station. This monitor information is forwarded to the monitor and control system and made available in the real-time database of that system and used for the station calibration algorithms.

In parallel with the filtering and beamforming the raw data can be stored in a transient buffer. Freezing of the buffer content can be controlled by internal or external triggers. The stored data or selections thereof can be sent to the central processing facility for further processing.

Each subsystem in the remote station is controlled by the local control unit, which will set parameters in the digital board driver. Typical control parameters are weights of the analog and digital beamformers, the clock rate and the filters selection.

The station synchronisation within the remote station provides the subsystems with clock and synchronization signals. For all receivers a coherent A/D converter clock is generated based on a local Rubidium signal. Additionally the synchronization between stations is done by disciplining the rubidium clock with GPS.

3.3 Wide area network

The station fields are distributed over a region of roughly hundred kilometres in diameter for the initial installation and potentially up to thousands of kilometres. For coherent data processing such as calculation of the correlation function and tied array beamforming, it is necessary to combine signals from all stations, which is performed at the central processing site. The wide area network transports the output data streams from the stations to the central processing site.

The main part of the LOFAR data that is transported throughout the network is observation data, in addition, a small part of the LOFAR data stream concerns monitoring and control data transport. The Wide Area Network consists of dark fiber and active communication equipment. Among these are network switches, optical multiplexers and media converters.

Data rates in the network vary from 2 to 20 gigabit per second, dependent on antenna location. The network will also be used for the transportation of data of the sensors of the LOFAR partners. The data from these sensors needs to be routed also over the LOFAR network. This has to be done in a way that other running experiments cannot be influenced.

3.4 Central Processor Facility

The main responsibility for CEP is to correlate the station data and deliver a data product which can be further processed by the user.

The Central Processing Facility is divided in three sections: an on-line section for processing of real-time data streams from the stations, a storage section collecting the processed data streams and making the resulting datasets available to the third section: the off-line processing. MAC is responsible to control CEP and allocates resources to particular observations, which can run in parallel with each other.

The on-line processing section consists of a Linux cluster for reception and pre-processing of the station data and transports the data to Blue Gene/L (BG/L) racks, where the data of the stations are correlated and subsequently reduced. Parallel to the BG/L resources, a general purpose cluster is available for auxiliary processing and in particular for real-time analysis, tuning and/or model creation tasks. The results of those tasks are typically used as control data for the processing applications running on the BG/L platform. The resulting data streams from the on-line processing sections are collected in the temporary storage subsystem.

The input section contains the connections to the stations (through the WAN). The data sent by the stations are sent in logical packages, each containing a time-frequency window of a single voltage beam. Each input node will receive data from the stations and run a data handling application that will buffer the input data and synchronise its output stream with the other input nodes based on the timestamps contained in the data.

Large amounts of processing power and internal interconnection bandwidth are provided through the BG/L supercomputer; a peak processing power of 34 TFlops is available for the processing tasks. This processing power in combination with the IO capabilities of BG/L allows for a correlator capable to handle 2926 baselines for the full 32 MHz bandwidth together with the channel filter.

The storage system provides disk space for the collection of data streams and storage of complete observation datasets for off-line processing. This storage is intended for temporary usage (typically 5 days) until the final data products are generated and archived or the raw data itself is exported or archived. Access to data in the storage system is through storage clients that have access to the metadata and file locations.

Finally a general purpose Linux cluster is used for the off-line processing. The off-line processing section offers general-purpose processing power and high bandwidth interconnections to the off-line processing applications. The largest part of this cluster is a “normal” Linux cluster computer optimised on cost per Flop.

The size of this section will be determined based on the applications performance, especially the self-calibration of imaging data for the EOR and surveys. Additional off-line processing power will be available in GRID networks, thus potentially providing extra processing power from remote sites. GRID networks also provide the basic infrastructure for data access and data export to user groups.

At the end of the pipeline, an archiving system will be available for long term storage or exporting the data e.g. to a Science Centre (Section 6.1).

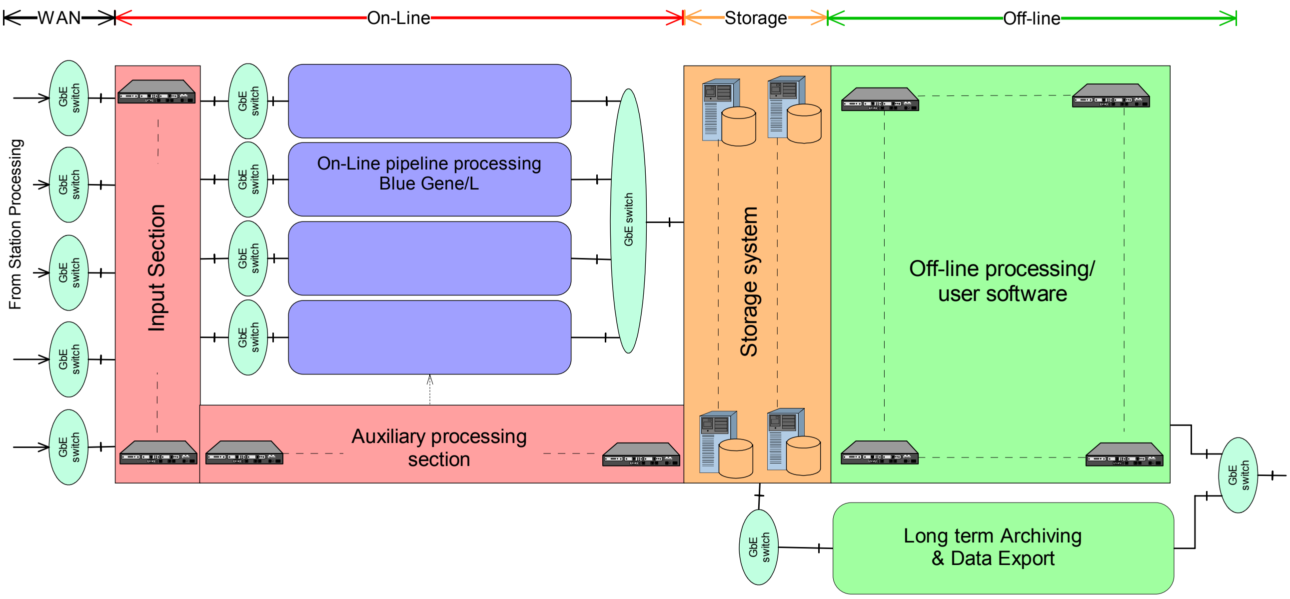

Figure 11 *Central Processor Facility hardware architecture overview. Four Blue Gene/L racks are embedded in a hybrid cluster computer. The sizes of the various parts of the cluster are scalable.*

3.5 SAS & MAC

In this section a description of the software systems with a focus on the control over the data processing pipelines is given. The *Specification, Administration and Scheduling* (SAS) and *Monitoring and Control* (MAC) packages are used to configure and execute the data processing pipelines. These pipelines can be described as steps in the production process of the final products of LOFAR. The final data products are sets of files which typically contain images, source lists etc. along with files containing metadata describing the circumstances during the observation and operational data from the data processing steps.

The specification and control related functionalities relevant in the data processing pipelines are described below:

- **Control room interface – prepare**
  The SAS GUIs are used to prepare observations. The user (scientist) uses these interfaces to propose[3] and specify observations. The observation specification is further translated in detailed observation and configuration settings by an instrument scientist, also using the SAS GUI and so-called observation tree definition templates. These templates contain the default values used for instruments or applications. These values may be pre defined by the developers or copied from

[3] Actually, the proposals are generated in a separate application called NorthStar. The output of that application is integrated into the SAS specification GUI.

previous observations. The result of these *specification* activities is a fully configured observation and instrument definitions held in the SAS Observation Tree DataBase (OTDB). Within the preparation three steps can be distinguished:

- **User Specification of observation**
  Observation specifications are entered in an *observation type* specific user interface. The user (scientist) enters observation specifications in terms of the required end product; these are further processed into settings for the relevant components in the observation. The generation of those settings is based on intelligent code in observation tree templates.
- **Detailed specification of observation**
  Based on the user specification most settings in the observation definition tree can already be entered. Some settings may not be generated automatically, or may be overwritten manually. These actions are performed by the instrument scientist, using the detailed tree configuration interface. For most observations this work should be minimal or not needed at all; this manual configuration is expected to be necessary only for special observations.
- **Configuration of the instrument**
  Based on the observation configuration and instrument configuration templates, the instrument settings are generated and further specified by the instrument scientist.

- **Observation planning**
  An automated scheduler is used to generate the observation planning, which is a sorted list of the observations with given start and stop times. The schedule generation is based on the list of specified observations, observation constraints, and instrument (resource) constraints. Additional manual scheduling can be applied by the instrument scientist. The scheduled observations are stored in the observation tree database (OTDB).
- **Observation execution and control**
  The execution of observation processes and data processing steps is performed by the monitoring and Control (MAC) system. The MAC system operation is based on the scheduled observations in the OTDB. Based on the configured instrument and observation initialisation and start times, the MAC system instantiates control trees for the observation and instrument control. These trees contain active controllers issuing commands to the (device drivers of) hardware components in the instrument and to processing tasks running on computers.
- **Observation monitoring**
  The observation and instrument control trees also monitor the status, progress and performance of the instrument and processing steps. The resulting monitoring values are stored in the MAC internal on-line database (PVSS).
- **System health monitoring and maintenance**
  The status of the instrument hardware parts is monitored continuously. This information is further enriched with sampled data from the observed data that are processed into metrics. Both the hardware status and metric data are input to a model based evaluation of the system health. This model reveals the resource status, malfunctioning and suspicious components. This information is stored in the real-time resource database that is used by the scheduler.
- **Control room interface – monitor the observational process**
  The monitored instrument status and metrics from the processing pipelines are available in the real-time database of the MAC system. These values are visible in the operation inspection panes of the MAC navigator. Multiple views on the instrument are available along with observation specific views. The real-time database is also programmed to generate events if the monitored values exceed pre-defined boundaries. Those events can be visualised with e.g. flashing red items for critical situations and may also result in e.g. automated messages being sent to cellular phones of operator or maintenance personnel.
- **Control room interface – inspection of data products and quality**
  After completion of the observation, and after completion of each processing pipeline step, the metadata can be inspected in the SAS inspection GUI. Along with the metadata the user and instrument scientist can also assess the observation quality by inspection of snapshot images and similar inspection products generated by the processing pipelines.
- **Data products dispatch**

The archiving and export of data products can be defined during the observation specification phase (see above); the specified actions will be scheduled and executed by the MAC system, similar to the execution of processing tasks. This will result for instance in a data transfer to GRID connected storage facilities. The same data dispatch actions can also be initiated interactively from the SAS observation inspection GUI.

The software architectural model concepts described above are input to the analysis model shown in Figure 12. This model shows how the various model elements are related to each other. The actual implementation of this architecture in LOFAR closely follows this model; this is indicated by the colours. The model elements define the main requirements in the MAC and SAS subsystems. Some of these functions are mapped directly on classes or databases in the implementation, such as the SAS specification GUI (OTB) and SAS databases. The MAC control functions must be seen as the highest level functional description of the tasks running on the CCU and LCU machines.
The choice for this model is based on the following considerations:

- The MAC system is in full control of all operations in the LOFAR system. Also, all systems are controlled in a similar way. Only at the lowest level are there specific layers to control software applications, signal processing equipment, infrastructure etc. This uniform control structure is accurately modelled by the three processing stacks in the lower half of the figure.
- All user interaction is through the SAS and MAC systems. MAC is used to interface to running observations or processing pipelines. SAS is used for all other interaction prior to execution and afterwards. There is no “direct” interaction with applications; application specific interfaces can be plugged into the MAC and SAS GUIs instead.
- There is separation of concerns between the SAS and MAC systems with clear interfaces (mainly database tables and configuration files). This way the control systems can be developed and tested independently of each other. Also during operation the MAC subsystem can operate while the SAS subsystem is down.
- The data processing flow is apparent in the model elements in lower half of Figure 12. Therefore, the data flow is also explicitly addressed in the controllers. This part of the architecture is implemented by hierarchies of control nodes. The lowest nodes in this hierarchy control the individual pipelines, while the higher nodes control the sequential behaviour between these pipelines. This hierarchy is continued also within the design of the processing pipelines control in which case the leaf nodes control individual executables in a pipeline. This representation of the data flow in the architectural model makes the control system easily extendible with new observation modes or even new subsystems.

ASTRON

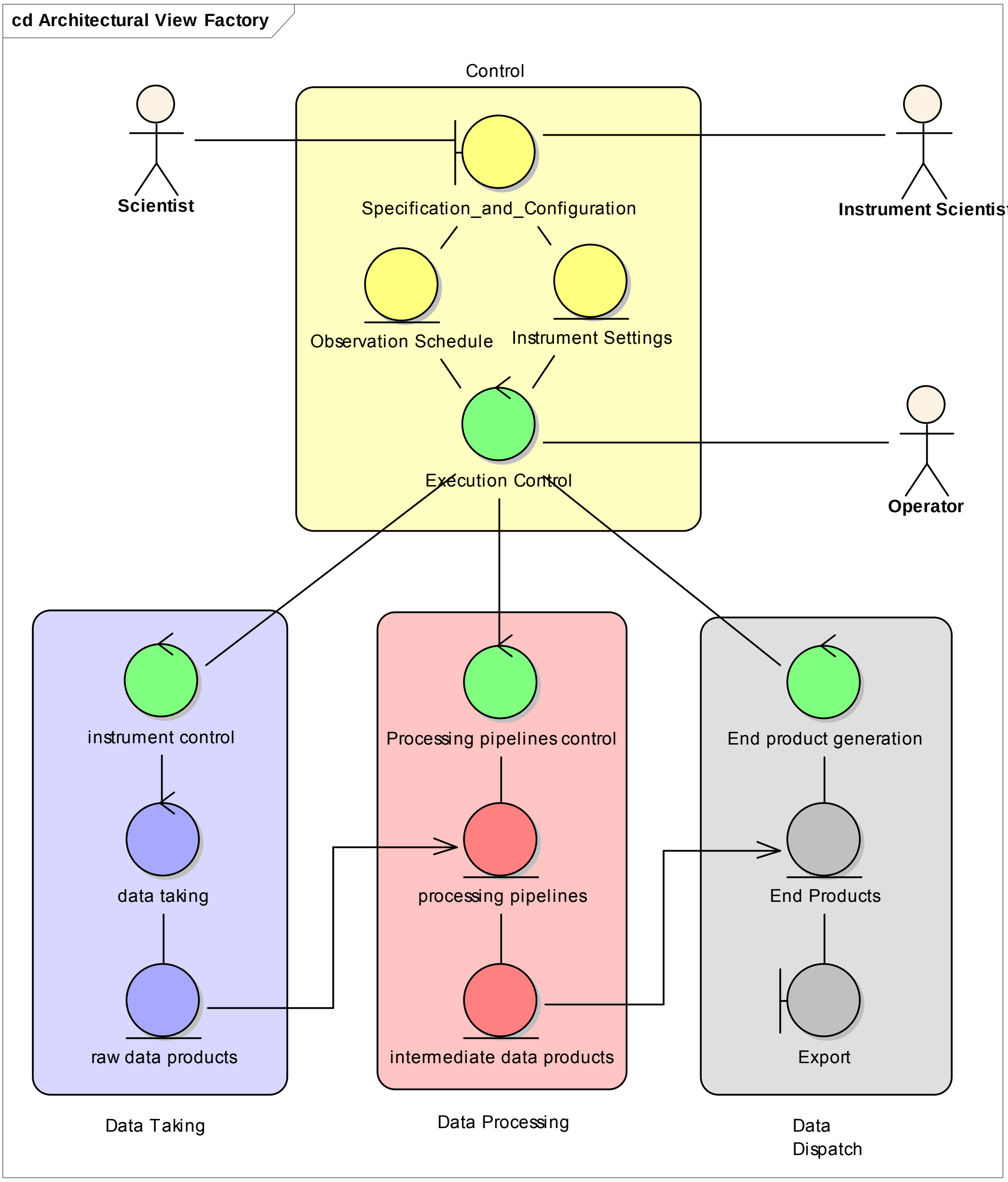

Figure 12 *Analysis classes in the data processing architectural view. The whole data taking, processing and end product generation process is controlled by the SAS and MAC components in the upper half of the diagram. Those components control the actual production process represented by the lowest two rows of classes. This production process runs from left to right and is represented by the data taking process (in blue), the data processing pipelines (in red) and the data dispatch (shown in grey). Note that the SAS and MAC analysis model components (shown in yellow and green respectively) correspond to actual components in the detailed design and realised implementation of those packages.*

3.6 System health management

The percentage of time during which the LOFAR system is effectively operational i.e. the system uptime is an important issue that warrants a lot of attention. The reason for this is that due to the complexity of the LOFAR system and the harsh operating environment, it is almost certain that at any moment in time many of LOFAR's components (antennas, amplifiers, network links, computing nodes, etc.) will be non-functional. Within reasonable bounds, this should not impact the usability of LOFAR for performing useful scientific measurements; rather, the system performance should gracefully degrade with each failing component.

Any faulty component may affect the quality of the measurements in a negative way, and may also jeopardize the operational capabilities of the LOFAR network. The objective of the SHM module of the LOFAR system is to support the efforts to maximize the system uptime. The main functions of the module will be the early detection of system failure, the accurate identification of failing components, and the support for remedial actions.

Daily or weekly on-site inspections cannot be performed in an economically viable way (at least for the remote stations). Hence the System Health Management subsystem will primarily be guided by the data that is generated by the LOFAR system in an automated fashion. This data consist of both the scientific data (generated by the antennas) and the 'housekeeping' data of the equipment that controls the sensor network. Deviations in system health will be reflected in sensor data that deviate from the normative measurements. These deviations are called symptoms and are used by the SHM module to detect system failure and identify the responsible system component.

4 Instrument geometry

The positioning of receptors in the array affects many aspects of the LOFAR design including signal transport, signal processing, calibration performance, site of the array and scientific goals. Optimizing the array geometry will, therefore, be subject to a number of 'hard constraints' from LOFAR sub-systems, science requirements and of course availability of land.

Limitations on data transport rates require that receptors located far from the central processing facility be grouped into stations where individual antenna signals are combined and transported at reduced data rates. Each station will comprise of order 100 receptors, and beams formed at the stations will reduce the total bandwidth requirement on long distance connections to the central processor. For receptors near the central concentrator node (where all data is coming together), this constraint is relaxed. Within a diameter of 2 km signals from each station have a bandwidth which is 10 times higher than the remote stations. This inner region is referred to as the core.

4.1 The sensor site

One of the main constraints of the funding agencies was that LOFAR should be built in the Netherlands. Since, the aim of LOFAR is to receive weak signals from the sky the RFI environment is crucial. Another requirement is the availability of land. Both arguments have led to centre LOFAR around Exloo in Drenthe. A number of RFI monitoring campaigns were set up and the results can be found in [19].

4.2 The LOFAR configuration

A total of 77 sensor fields are distributed along 5 log-spiral arms with a diameter of approximately 100 km. This is shown in Figure 13.

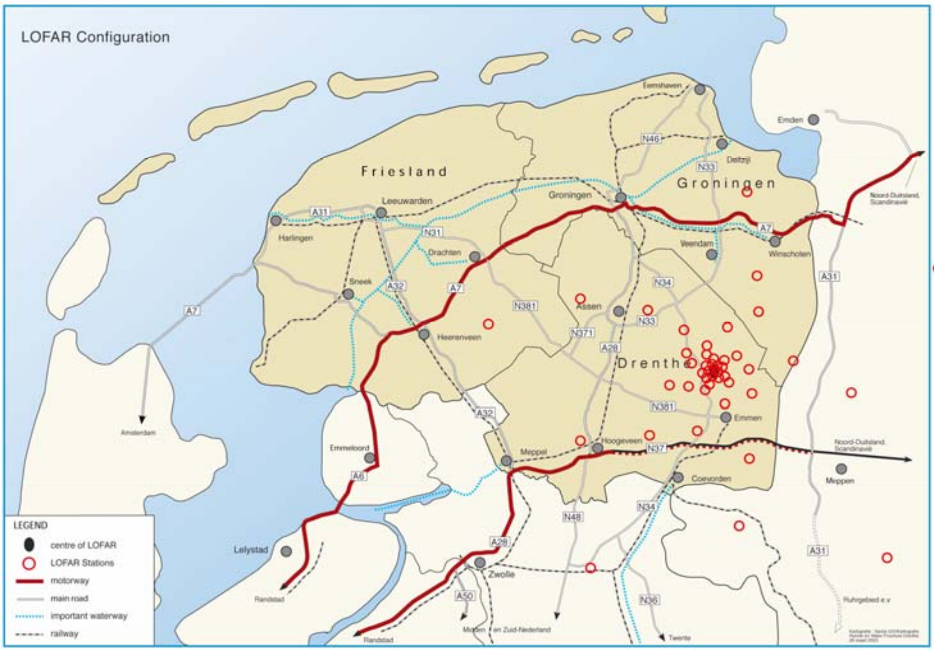

Figure 13 *Detailed distribution of LOFAR stations over the north of the Netherlands.*

The sensor fields are centrally condensed, following a logarithmic distribution, resulting in an inner area of 2 km diameter where about 40% of the stations are located (see Figure 14). This inner area (core) can be operated for dedicated experiments yielding more data per station than what is achievable with the “outer” stations [1]. This makes the core suitable for “all-sky” monitoring programs. The core can also be used to calibrate the large ionospheric phase fluctuation that would otherwise lead to severe decorrelation when correlating remote stations. The adopted calibration scheme does not depend on this approach, but the core has sufficient sensitivity to leverage sensitivity of the much smaller remote stations for the calibration of the ionospheric phase screen.

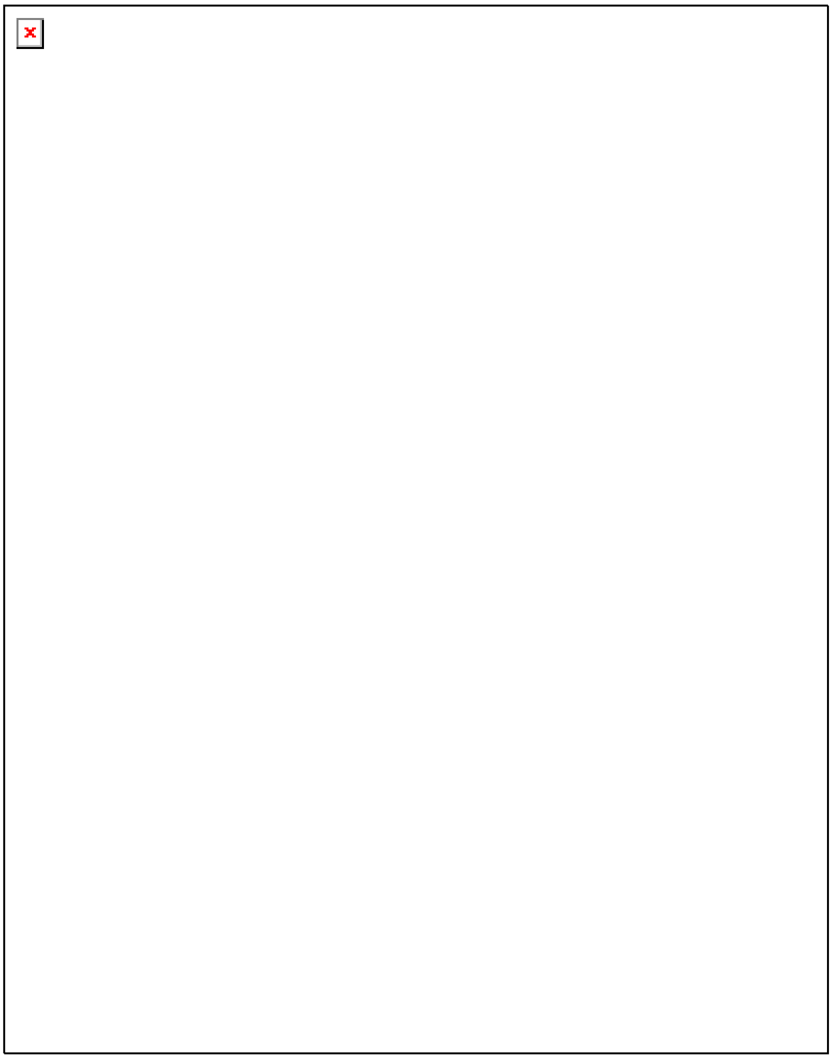

Figure 14 *Distribution of station fields in the core of LOFAR. Shown are the stations located in the inner 2 km of the array. The orange circles are the low band antenna arrays and the yellow octagons are the high band antenna fields. The light grey area is a historical canal.*

The configuration will be two-dimensional in order to yield good instantaneous UV coverage. Furthermore, the configuration is a balance between imaging capabilities under a variety of observing conditions and mainly infrastructural cost. Infrastructural cost is the main reason to define the arms in LOFAR. Furthermore, an odd number of arms are chosen to break symmetry as much as possible. Taking the rotation of the earth into account the UV coverage can be simulated. This was done for a number of configurations [2] and led to the current configuration. Also the availability of land was taken into account.

4.3 The station configuration

The placement of receptors in stations determines the effective collecting area over the LOFAR observing band. The optimal spacing of antennas for a given frequency is at half a wavelength, maximizing the station field of view for a given collecting area and minimizing power in the station beam sidelobes. The wide frequency range to be covered by LOFAR antennas forces compromises, however. By choosing different spacings over the area of the station and by applying tapering functions in the beamforming one can optimize the collecting area as a function of frequency, while grating lobes are smeared out. The number of receptors and specific layout also determines the beam characteristics. LOFAR calibration will solve for these station beams.

The sensor fields contain two arrays of 96 antennas [3] each distributed over an area of ~100 m diameter. The low frequency band of LOFAR (30-80 MHz) is sensed with dual polarized wire dipoles, while the high frequency range (120 – 240 MHz) is sensed with 4x4 phased arrays containing dual polarized bow tie dipoles. For the antenna configuration of the LBA an irregular array is chosen to scramble the grating lobes. A more detailed motivation can be found in [10]. Electro-Magnetic simulations of different configurations are presented in [20], [21], [22], [23], [24]. The irregularity is achieved by using an expo shell configuration as the basis, with on top of that a perturbation on all the element positions with a random offset which equals 10% of the radius of the shell where that element is located. An example of such a configuration is shown in Figure 15. The goal is to generate a unique randomized configuration for each station in order to scramble the grating lobes from station to station in a different way. An alternative for this is to rotate the station configuration with respect to each other.

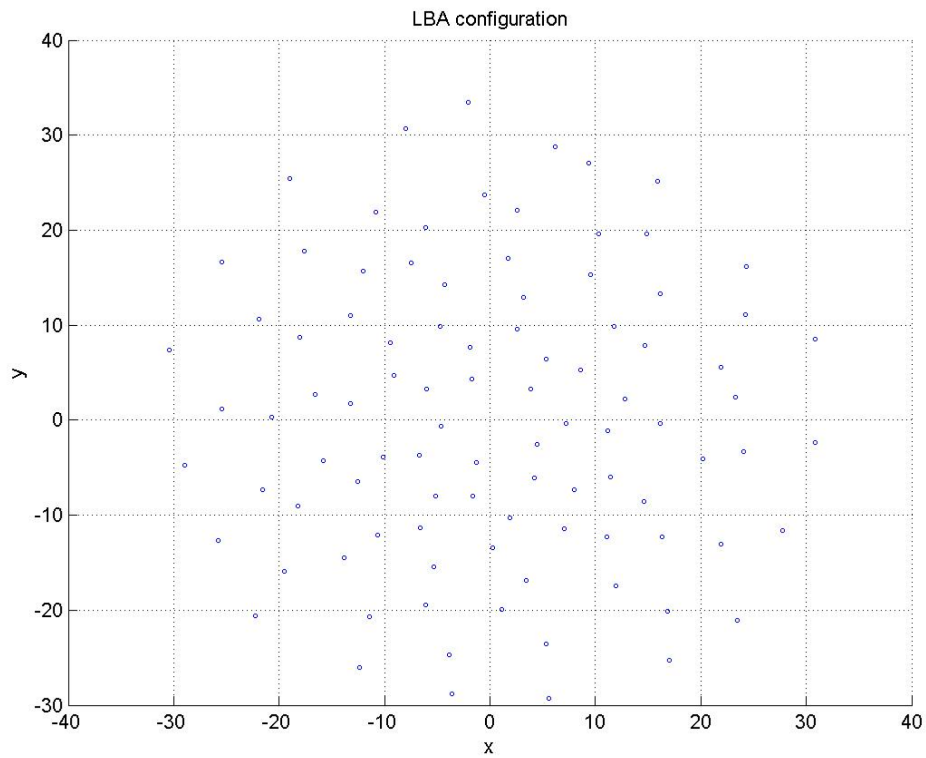

Figure 15 *LBA configuration of one of the stations currently in the field (axes scales are in meters)*

The High Band Antenna (HBA) configuration is not yet finalized. However the 4x4 array, further referred to as tile, will have elements which are equidistantly spaced at 1.25 m to maximize the collecting area. It is expected that the grating lobes can be dealt with in the calibration or that they can be smeared out by randomizing the HBA station configuration or rotating the HBA station configuration with respect to each other.

The LBA and HBA antenna fields are placed in a separate area within the station footprint. Both will be arranged at different sides of the station electronics to save on cable length and economical use of land.

4.4 The central systems site

The sensor fields produce data streams which are transported to the central area in Exloo (the concentrator node) over optical fibre links. From there, all data is sent to the computer facilities in Groningen for further processing and storage of the data. The RUG is one of the LOFAR partners and is experienced in, and has facilities for housing, large clusters of computers. Transporting the data first to Exloo gives the flexibility of installing processing power within the core that allows housing local processing power which may be used to develop alternative EOR, tied array or transient modes.

5 Calibration and RFI Mitigation

5.1 Calibration approach

The development of a robust calibration system for LOFAR is among the most challenging and high-risk tasks within the project. The calibration will take place in two main stages: first in each station, further referred to as the station calibration [25] and second in the central systems [40]. The goal of the station calibration is to calibrate individual antenna signals for gain and phase differences. These differences are caused mainly by temperature variations in the electronics. In the central systems the differences between stations will be calibrated using an approach that resembles conventional self-calibration in synthesis arrays. Depending on the scientific application this will be done in (almost) real time or off-line at the central processing site within about a week. The Calibration Framework has been described in detail in [40]. The current status of the process to validate this framework is described in [41]. Below we will briefly describe some essential concepts, the differences between conventional and LOFAR calibration, and the main challenges faced within LOFAR calibration.

5.1.1 Conventional self-calibration

Conventional self-calibration, which has enjoyed spectacular success over the last quarter-century, solves for a single complex gain per antenna for each time interval. A model of the sky brightness, combined with the current best estimate of the complex gain of each antenna (including the phase delay due to atmospheric propagation of the signal) is used to predict the visibility values. These predicted values are then compared to measured values, and an analysis of the differences leads to a set of adjustments to the antenna-based gains, which minimize that difference. The resulting data set is then imaged in order to create an improved model of the sky brightness. The non-linear image deconvolution step imposes generally justifiable constraints on the model brightness distribution, and is essential for translating improved calibration into an improved sky model. The loop is iterated to convergence.

5.1.2 LOFAR generalized self-calibration

Fundamentally, LOFAR calibration is position-dependent. The station beams change strongly with time as earth-rotation alters the projected geometry of the fixed dipole antenna arrays. Each station will be laid out differently as is discussed in Section 4.3. The ionospheric delay as viewed from a given station is strongly variable as a function of position within the field of view. The low-frequency sky is sufficiently well filled that position-dependence of the gains cannot be ignored. A more sophisticated approach than conventional self-calibration is therefore mandated.

A key point is that it is not possible to take an uncalibrated LOFAR dataset, apply station-based multiplicative corrections, and arrive at a fully calibrated dataset suitable for high fidelity imaging. Such corrections can be valid only for a single point on the sky, implying that the resulting dataset is miscalibrated for all other positions on the sky. The resulting sidelobe responses will result in unacceptable imaging performance. Instead, to get a dataset that can be imaged with high fidelity, one must evaluate the complex gain that applies to each source, subtract the contributions to the visibility for each source, then add them all back in with unit gain. Therefore, we take the dataset with the sky model subtracted source-by-source, and complex gain by complex gain, then image the residuals. Such residual imaging must be performed by "tiling" the field of view with many small patches. For each patch, the residual data are corrected for a single sky location using the calibration parameters, and the rest of the sky can be considered to be empty to an acceptable level. To the resulting residual map, we can then simply add the sky model to arrive at a high fidelity sky image.

Ideally, a generalized self-calibration system would produce an independent complex gain for each point on the sky, for each station, and for each time interval. This set of complex gain estimates would be used to subtract out the sky model, point by point, yielding visibility residuals which can be analyzed to refine estimates of the complex gains. Errors in the sky model would be estimated by imaging of the residual data set. Conceptually, this is closely analogous to conventional self-calibration, except that the number of complex gain parameters to be solved for is larger by a factor corresponding to the number of independent locations in the field of view. There are, however, two major practical difficulties.

- If all times and all sky positions are independent, the total number of parameters to be solved for is excessive, and the robustness of any algorithm will be compromised.
- Nonzero values in the residual dataset result from the sum of errors from sources in all directions. It is not trivial to devise methods for analyzing the residuals to derive a complex gain estimate in a particular direction.

The LOFAR calibration system, and indeed the whole LOFAR instrument, is being designed with these two problems firmly in mind.

5.1.3 Solutions to LOFAR calibration challenges

Number of Parameters

In essence there are three main types of parameters in LOFAR calibration: the sky, the beam and the ionosphere (most of which also depend on direction, frequency and time). The total number of parameters that need to be solved for may therefore become prohibitive. The solution may, in fact, not be robust. Or it may converge only very slowly, if at all. It is therefore explicitly recognized that most calibration parameters are smooth functions of position, frequency and time, and so the total number of independent parameters required is reduced dramatically. Solutions in the various parameter domains are then computed in the form of polynomial coefficients on various scales. LOFAR has been designed in such a way that the sidelobes of the synthesized beam are low (achieved by using a relatively large number of stations and using wideband synthesis) while a station is designed such that the beam sidelobe suppression will reduce the number of distant sources that need to be taken into account (achieved by a randomized station configuration and a relative station rotation).

During the startup phase of LOFAR a Global calibrator Sky Model (GSM) will be built up using external surveys combined with a shallow survey over a wide range of frequencies using LOFAR itself. Once LOFAR is in full operation the (bright source) sky will become progressively better known and the computational effort required to generate source parameters will be sharply reduced.

Calibrator Source Spatial Isolation

The second problem (residuals have contributions from the entire field of view) demands a way to effectively isolate a small patch of the sky by suppressing contributions to the visibility from outside that patch. If such a spatial filter can be applied to the dataset, the resulting filtered dataset can be treated with a single complex gain per antenna. By rotating a dataset to a chosen phase centre, and performing time and frequency averaging, it is possible to restrict the field of view in this way. Typically, the patch of sky being isolated will be dominated by a strong point source. This source is subtracted using the best estimates of the calibration parameters derived from the filtered dataset, and the procedure moves on to the next brightest source. The process of comparing actual and predicted visibilities and deriving calibration parameters takes place one source at a time, instead of all at once. This progressive subtraction process has been named “peeling”, as in the layers of an onion.
A crucial fact regarding the calibration system architecture is that there are many sources within each field of view that are bright enough to be individually isolated and solved for in a closed-loop process. Provided that the solutions in the direction of these sources are not significantly contaminated by nearby sources leaking in through imperfect spatial filters, such closed-loop calibration can be expected to reach noise-limited performance, as is routinely achieved in conventional self-calibration.

Remaining Issues

Questions remain as to the destabilizing effects of imperfect spatial filtering, and as to the sensitivity of the generalized self-calibration process to large fluctuations in the effective number of calibration parameters as a function of observing parameters and conditions. The total processing power required to perform the above functions is also uncertain, and depends on the speed of convergence of the major self-calibration cycle. The signal to noise ratio that is achievable on calibration sources depends on the array design, and upon ionospheric phase coherence times. There will be ranges of conditions and observing parameters that render LOFAR essentially uncalibratable. While current estimates lead to the conclusion that the excluded observational parameter space is reassuringly small, uncertainties remain. In particular, the properties of the sky at low frequencies and high resolution have strongly influence on the calibratability of long baseline data, and require improved understanding.

5.2 RFI Mitigation Approach

LOFAR is among the first radio telescopes for which RFI mitigation forms an integral part of the system design. This is needed, as LOFAR will operate in a hostile RFI environment. Unlike the majority of present-day radio telescopes, the signals in most of the frequency bands at the input of the LOFAR A/D converters can no longer be characterized predominantly as white noise. RFI mitigation in radio astronomy is a relative new field, and many RFI mitigation methods are not yet "fully proven". However, many techniques have been extensively demonstrated on prototype systems. The overall LOFAR approach towards RFI mitigation is to rely in the first place on such well-proven techniques, but to provide also options to study new and promising approaches. Their application will offer enhanced performance in contaminated bands, and ensures that LOFAR can handle the challenges imposed by future changes in the use of the radio spectrum. It is therefore vital to ensure that the LOFAR system is designed in such a way that future, more complex RFI mitigation algorithms can be applied without redesigning parts of LOFAR. LOFAR also offers a unique opportunity to develop and demonstrate new techniques for application in future instruments, in particular the Square Kilometre Array.

A major concern related to RFI is the way in which the interfering signals propagate through the LOFAR system, finally ending up as a distortion of the measurement results. Issues here are for example the attenuation due to fringe rotation, delay smearing, or beamforming sidelobes. Also, the effects of the proposed mitigation methods on the data are insufficiently known: both the weak interference residuals after RFI mitigation and the RFI mitigation processing itself may distort the astronomical data.

The impact of RFI mitigation on the system design can be roughly categorized as follows:

- The *linearity* requirements on the analog systems (antenna Low Noise Amplifier, receiver unit) and the *number of bits* in the analog to digital conversion and the early phases of digital processing are largely determined by the strongest signals present in and outside the bands of interest.
- The *location* of antenna fields relative to transmitters determines the strongest signals at the location. Also, the distance to the nearest power lines, roads and villages should be taken into account as well as pulsed emissions from electrical fences.
- The *architecture* and *processing budgets* for all digital processing subsystems is influenced by the required active RFI mitigation techniques. Relevant issues here are for example computational power (especially for correlation and matrix operations such as inversion), transfer speed, and data size of parameters (relevant for RFI mitigation with feedback loops).

The strategy consists of creating several "lines of defence" with increasing complexity, and therefore increasing risk. Simple and proven techniques are expected to be sufficient in the majority of environmental conditions. When necessary, more complex techniques can be applied. While these techniques may require more experimentation in the early phases of operation, they will lead to increased system performance as they become more mature. The adopted strategy assumes that (a) the analog subsystems and digital filters are sufficiently linear and (b) low level semi-stationary RFI sources can be removed from visibility data and images through self-calibration and other post-correlation techniques. Results from prototype antenna and receiver systems indicate that the first assumption can be met [30]. The second assumption has not yet been sufficiently validated. The generalized self-calibration process is in principle capable of handling low-level RFI, but no definite proof through simulations or processing of realistic data has been provided yet. It should be noted though that many post-correlation RFI mitigation approaches have been tested at existing synthesis telescopes. Post-correlation mitigation has the major advantage that information from all stations is available. This leads to a natural spatial suppression of localized RFI (also through spectral dilution). It often also allows for a better parameterisation (and suppression) of RFI than possible for stations separately because of the spatial properties mentioned above, but also because of the relatively long integration times at post correlation level.

5.2.1 Mitigation strategy

The first step in the LOFAR RFI Mitigation strategy [30], [31] is to select the cleanest parts of the spectrum by choosing the optimum location of the subbands[4] in the available bandwidth. In a large fraction of the LOFAR band the time-frequency occupancy is less than 10% at the level of 5 dB above the sky noise [36] for a single antenna. The majority of the RFI sources have a bandwidth of order kHz, so it is possible to avoid them without discarding more than 10% of the available signal bandwidth. At lower power flux levels, the time-frequency occupancy of transmissions may be different, but several 12 hour duration Low Frequency Front End (LFFE) observations at the Westerbork Synthesis Radio Telescope (WSRT) [38] indicate that 80% of the selected best 8 x 2.5 MHz wide bands in the 110-180 MHz range were usable. The loss of sensitivity in those observations therefore was limited, and the sensitivity in the 80% best part of the bands was sidelobe and confusion limited. Although not solid proof, this experience indicates that especially for (non-line) observations with processed bandwidths that are significantly less than the maximal bandwidth, LOFAR observations can be carried out with sensitivities at least as good as with the WSRT-LFFE.

The frequency band selection process relies on determination of RFI levels right after the filter section in the station digital processing [33] and also after the channel selection filter at the central level. Relatively simple statistical properties (power detection [34] followed by estimating max, median, and percentile values) are sufficient to detect the RFI signals above sky noise.

The RFI in the remaining 80-90% of the band will typically be less than ~ -15 dBμVm^{-1} at 1 kHz channel bandwidth or ~ –190 dBWm^{-2}Hz^{-1}. Station beamforming will attenuate these sources to the level of strong sky sources [31]. This allows for further mitigation in the post-correlation domain, in particular as part of the self-calibration process.

When more than ~20% to 40% of the band is affected by RFI, an option is to apply spatial filters at subbands and/or at the kHz post correlation level. Most of the strong interferers are located at fixed positions at the horizon, allowing for effectively attenuation through deterministic nulling. Nulling fixed directions is a well-conditioned process, avoiding the need for real-time estimation of the Direction Of Arrival (DOA). Beamshapes vary smoothly while tracking celestial sources. Only where strong moving interferes are present, adaptive nulling techniques have to be employed, but only if the sources move gradually such that the beam sidelobes change smoothly over time. Several DOA estimation techniques have been extensively studied through simulations. LOFAR will be an excellent platform to further develop these spatial filtering techniques for astronomical applications. Artifacts induced by spatial filtering have been successfully suppressed in some experiments [34], [35].

5.2.2 Mitigation in relation to the LOFAR architecture

In relation to the LOFAR architecture, RFI detection and mitigation can be dealt with in the field (at the station locations) or centrally [32]. This depends on where certain information is, and on signal properties such as time-resolution, frequency-resolution, and spatial information content of the signals under consideration. It also depends on required and available computational resources for the mitigation processing steps. Since interference mitigation is a relatively new field for radio astronomy, experiments with the real system are necessary to judge where in the signal chain the final algorithms should run. It also

[4] subband: the 156 kHz / 195 kHz wide frequency bands at station level, channel: the 1 kHz wide frequency bands at central level

should be possible that all RFI mitigation algorithms are switched off. This will especially be the case during initial operations. The design should be suitable to accommodate RFI detection and mitigation algorithms at several stages in the system.

Table 2 lists the possible mitigation options and their intended application in LOFAR. A detailed explanation of these techniques, especially detection and spatial filtering, can for example be found in [34].

Table 2 *Summary of mitigation options and their intended application*

Where		Technique	Intended usage
Station	Before sub-band filtering	Fast blanking of raw time samples	No: breaks time-invariance, unknown influence on interferometric gain
Station	Before beam-forming	Selection of RFI free subbands	Yes, requires detection and short/long term statistics, both for subbands (station) and channels (central)
Station	During beam-forming	Fixed direction spatial filtering (subbands)	Yes, only if needed; no. of directions << no. of antennas
Station	During beam-forming	Adaptive spatial filtering (subbands)	Optional, experimental, drawback: hampering of the calibration process
Station	After beamforming	Blanking of subband signals	No: breaks time-invariance
Station	Before or after subband filtering	Parametric RFI removal	Optional: complex, experimental
Station	TBB	Detection, blanking, spatial filtering, parametric methods	Yes, if needed, experimental
Central	Pre correlation	Fixed spatial filtering (on subbands or channels)	No, spatial filters will be applied at post correlation level (channels)
Central	Pre correlation	Adaptive spatial filtering (on subbands or channels)	No / optional: complex, experimental
Central	Pre correlation	Pre-correlation parametric RFI mitigation (subbands or channels)	Optional: complex, experimental
Central	Post correlation	Post-correlation blanking / flagging (1-100ms, 1kHz)	Yes
Central	Post correlation	Post-correlation spatial filtering (1ms-60s, 1kHz)	Yes, experimental
Central	Post correlation	Post-correlation parametric RFI mitigation	Optional: complex, experimental
Central	Post correlation	Self-calibration processing	Yes (reducing to noise level)
Central	Post correlation	Correction of distortions and artifacts	Yes, but experimental

At remote station level, and for the antennas in the core, the power in each subband will be calculated in order to detect RFI in the spectral domain [33]. RFI detection will be quantified by power levels relative to threshold values as a function of frequency. The thresholds are obtained during station calibration measurements. Both time continuous and intermittent RFI sources can be detected. Based on these detections (percentile statistics), optimal subbands can be selected. This process will be refined by applying similar statistics at kHz level at the central site. The detection intervals are of order 100 ms, and in most cases match the observed and expected time variations of the occurrence of RFI signals.

When needed, deterministic nulling can be switched on in the station beamformer, attenuating signals from interferers at fixed locations at the horizon by adjusting the complex weights for each spectral subband. The location of the interferers will be determined either through calibration measurements or through direction of arrival algorithms. Additional RFI detection processes may refine spectral detection by comparing output beams for successive frequency channels. For all applied interference mitigation methods it is crucial to log the RFI mitigation parameters for later usage during analysis.

Optionally, adaptive spatial filtering can be used at stations to put nulls on moving interferers. This does not add computational complexity to the beamformer. However, the DOA algorithms require additional processing resources. Some algorithms require calculating the covariance matrix of the incoming antenna signals and require matrix inversion, which adds significantly to the required computational power. Other schemes employ reference beams to monitor and characterize RFI sources continuously. As the LOFAR calibration requires subtraction of known strong astronomical sources, only smoothly changing nulls are acceptable. This is because otherwise the astronomical source subtraction process will fail because of unwanted changes in the spatial sidelobes.

The above techniques can in principle also be applied at the central site before correlation, using either subbands or channels. Because of limited available computational resources at that point, and because these techniques need to be applied on-line (at that point), applying flexible off-line post-correlation spatial techniques is the preferred option, at least in the earlier stages of LOFAR.

Flagging (blanking. excising) is a well-known technique in radio astronomy and will be applied in LOFAR at post-correlation level. This technique will not be applied at station level as it has several severe drawbacks when applied at the stations (e.g. it does not preserve the time series coming out of a station).

Parametric techniques, especially those techniques which make use of explicit knowledge of the RFI modulation or coding schemes [36], are interesting options to suppress for example digital radio (DAB-T) and digital television (DVB-T). These techniques are experimental, and additional research is needed to verify their use for suppression of e.g. digitally coded signals.

The discussion above is mainly concerned with the imaging mode of LOFAR. For other modes, especially transient modes all interference mitigation options are open. As a large fraction of the signal processing for those modes is off-line, RFI mitigation can be applied off-line when needed.

6 Operations

6.1 Operations model

The LOFAR operational model is designed with flexibility in mind. The instrument itself can be controlled in great detail and adapted to the specific requirements for a given observation. Multi beaming, and therefore multiple concurrent observations, is at the heart of the specification and control systems. Fast switching between observation modes is supported by the scheduling and control systems as well in the data processing and distribution pipelines. This same concurrent usage paradigm is exploited to perform calibration and analysis tasks while new data is being observed.

The users of LOFAR will interact with the instrument over network connections, as is shown in Figure 16. In general, these users will be operators at the LOFAR operations centre, but might also include "expert" users located at the various Science Centres. All stages of the observation process, from preparation through execution, storage, and export, can be accessed over the network using graphical interfaces. These interfaces provide full control over all available functionality; therefore, there is no need to be physically located at any part of the instrument.

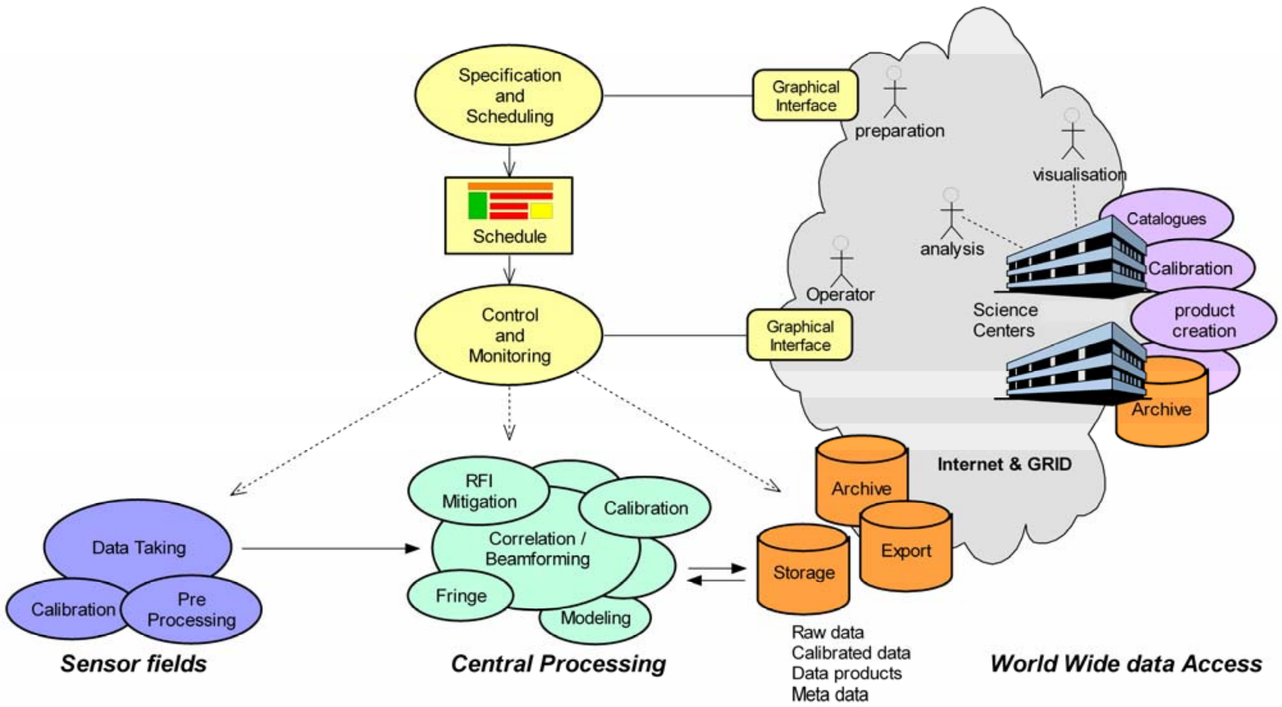

Figure 16 *Operational view on LOFAR. Shown is how the processing tasks at the bottom left of the figure are controlled through Specification, Scheduling, Control and Monitoring software. This management software is operated by the observatory staff and by the Science Centres when applicable. Detailed post-processing tailored to specific science needs may be done at the Science Centres.*

The operation of the standard LOFAR processing will be the responsibility of the LOFAR Operations Centre (LOC). Several specialized astronomical user groups will host Science Centres. These centres will have dedicated high bandwidth data connections to the central processing facility and provide data storage and processing capabilities to their users.

The instrument will be operated in fully automated mode as much as possible. All processing tasks are executed in pipelines operating in general on either real-time data or on datasets stored in the temporary storage facility. Data export and archiving are fully guided by the control software and will utilize GRID middleware to distribute LOFAR data products either to a permanent LOFAR archive or multiple archives located at the various Science Centres. None of these potential long term archives themselves are formally part of the LOFAR system.

GRID infrastructure plays an important role in the post-processing of data as well as in the distribution and archiving of data products. However, processing that is directly related to the LOFAR instrument itself, such as calibration, is typically performed at the central processing facility. In general, all data processing that requires huge data volumes and processing requirements is typically executed at the central site due to its tailored processing and storage capabilities. Remote resources, available through GRID interfaces, will be used for further processing of reduced datasets, certain reprocessing tasks, and permanent archiving, although this may change over time.

The architecture of the LOFAR instrument is explicitly optimized for flexibility in use and re-configurability of the hardware resources. In this way, many different types of astronomical observations are possible.
The clearest example of such flexible usage is the possibility to perform multiple observations concurrently. These observations may be of the same observational mode, in which case multiple beams may observe different part of the sky and can serve multiple users. Similarly, multiple observational modes can be operated simultaneously which will be especially useful for monitoring or piggy-back modes. For example, the same data stream being produced as part of an imaging mode can be queried continuously in real-time to search for transient sources. This flexibility is also dynamic, allowing the system to respond to internally or externally generated triggers and switch between observations and modes rapidly. Taken together LOFAR will be able to achieve a high observing efficiency.

Re-configurability of hardware resources allows for the definition of different observational modes, now and in the future. Most of the resources can be re-programmed, for instance from correlator to beamformer. The re-configurability is very clear in the central processing facility. There, a completely re-configurable data-handling framework is offered to the applications. The actual processing steps are partly executed on the same microprocessor systems, again offering re-configurability at the programming level. The central processor can be seen as a large collection of resources on which applications defined in software are executed.

The station system offers re-configurability at the operational level. The whole processing chain operates on maximal 215 independent subbands that can be configured into many different beams[5]. The same configuration freedom can be used to construct sub arrays within the station.

6.2 From proposal to user data product

For normal operation, human intervention in scheduled observations is not required. The LOFAR telescope is able to run in fully automated mode. Once observations are specified, the software system takes care of the scheduling process as well as execution of the observation data taking and processing steps. Standard LOFAR processing will consist of fully automated pipelines, typically encompassing task such as calibration, production and export of data products. Additional processing pipelines specific to a given scientific investigation will be available which execute on either the LOFAR Central Processing systems or utilizing additional resources provided by the various Science Centres.

The automated mode described above is what the system is designed to do and probably will do for the majority of observations, especially for survey and monitoring style observations and long campaigns. On the other hand, the software system should be flexible enough to allow for end-user and expert interventions in

[5] The processing power in the control systems currently limits this to 8 beams max.

all stages of the process from configuration to data export. So the above described automated system must have "open" interfaces and allow for extension or modifications at each stage of the process.

This section will describe how this operational behaviour is addressed in the main software subsystems constructing the observation process. More information can be found in the software top-level architectural description [43].

6.2.1 Proposal

The whole observation process starts with the observation proposal. The end-user constructs a LOFAR specific observation proposal using the NorthStar tool developed by ASTRON as part of the Radionet project. The Northstar tool contains specific modules for the LOFAR instrument. This tool delivers a complete observation project document containing the key observational specification parameters as well as the necessary information to prioritize observations based on scientific value. These configuration parameters represent the inputs for the specification and configuration step and are exported as XML files.

6.2.2 Specification and configuration

The observation parameters are the starting point of the specification and configuration process and are performed in the SAS system. Based on the key specification parameters an observation mode is selected. This observation mode corresponds to an observation tree template, which is a complete instantiation of an observation in the selected observation mode. The remainder of the specification and configuration process consists of filling this tree with parameters.

The observation tree contains all required parameter settings for the (automated) execution of the observation. This includes both hardware settings and application parameters. The majority of those parameters will have default values based on the selected mode. These default values are provided by the application developers and instrument engineers. Some parameters may also be deduced from the key observation specification parameters by scripts embedded in the SAS GUI.

The description of the specification and configuration process given so far describes the pre-defined operational modes. However, the SAS system also contains functionality to define new observation modes which may be variations or combinations of existing modes. The SAS Observation Tree Browser (OTB) Graphical User Interface (GUI) contains special panes for tree template construction. Once a new tree template has been made, it can be stored and used in a similar way as described above.

More detail about the specification and configuration process can be found in the SAS ADD [8].

6.2.3 Scheduling

The goal of the scheduler is to sort the list of specified observations in time. This sort operation has to take into account many aspects like:

- Expected performance of the observation in view of source visibility, sky interferers, earth interferers, ionosphere etc.
- Availability of instrument resources and constraints on combination of observational modes.
- Sequencing of sub-observations (measurements)
- Availability of supporting personnel, maintenance, activity planning etc.
- Observation priority, piggy-back modes, filler observations and fail-over observations

An automated scheduler engine will be built allowing human intervention. The human intervention can be used for all types of special constraints or wishes that are not easily modelled in constrain rules. A sort algorithm is used to calculate the optimal schedule given those human interventions and the programmed constrain rules [8].

The scheduling process is performed as a volume planner for periods of months, as an observation planner for week timescales and as a fine-grained planner on an hour to hour basis. This last planning level specifically takes into account the actual resource status of the instrument, actual ionosphere etc.

6.2.4 Observation execution; data taking and on-line processing

The MAC system takes care of the execution of the scheduled observations. While observations are being executed, the MAC system monitors the performance of both the instrument and the observations. This information is summarised in the resource database, which in turn is used by the scheduling system to update the schedule. This loop typically runs on a daily basis, but may also be used on short timescales to allow the system to adapt to situations where failures occur or the environment changes. This quick adaptive behaviour is a fundamental part of the system design. Beyond making the system fault-tolerant to hardware failures, it is currently undecided exactly which environmental or other system criteria may be useful and possible to monitor. Individual science applications may wish to monitor different system and environment characteristics of particular interest to their program. The MAC system design provides sufficient flexibility to support this option.

The MAC system controls everything that happens in the instrument, including the execution of the processing pipelines. The on-line processing pipelines in particular are executed in a completely analogous manner to the data-taking process and are fully automated. As discussed in the next section, the off-line processing pipelines are also controlled by MAC, but differ from the on-line processing in that they allow for the possibility of manual interaction during execution. This human interaction is possible through the use of the real-time database of the MAC system. This database is the interface between the MAC controllers and the device drivers or applications to be controlled. The operator can change settings in this database using the MAC navigator, thus directly influencing the observation process. However, most of such intervention is supposed to be programmed into the controllers (logical devices and instrument controllers). The nodes in the controller hierarchies contain rules on how to react to faults or other progress metrics. The MAC real-time database can also be programmed to generate alarms in the event of faults or if specified threshold tolerances are exceeded.

6.2.5 Off-line processing pipelines

The off-line processing pipelines operate on the data stored by the data taking process. For example, in the case of imaging applications, this data will consist of un-calibrated UV data. These pipelines are executed on the off-line processing cluster and are controlled by MAC in a similar way to the other on-line pipelines and hardware systems. The MAC control system includes special controllers to manage the logic of a pipeline (sequential execution, intermediate file storage, etc.).

From an operations viewpoint, a pipeline can be executed in a fully automated manner without user interaction. Such automated execution will be scheduled by SAS. It is also possible to operate a pipeline interactively. In this case, the MAC system will still control and schedule the resource usage, but the user or operator will have the option to manage the pipeline execution by interacting with the pipeline controller.
It may also be possible for the user to interact with the individual application controllers directly in order to evaluate the performance of any given application and potentially modify settings to adjust the process on the fly.

The main difference between automated and interactive operation is the origin of the control commands to the pipeline controller. In the automated mode, these commands originate from higher levels in the control tree. In interactive mode, the SAS and MAC GUIs drive the pipeline controller directly as illustrated in Figure 17.

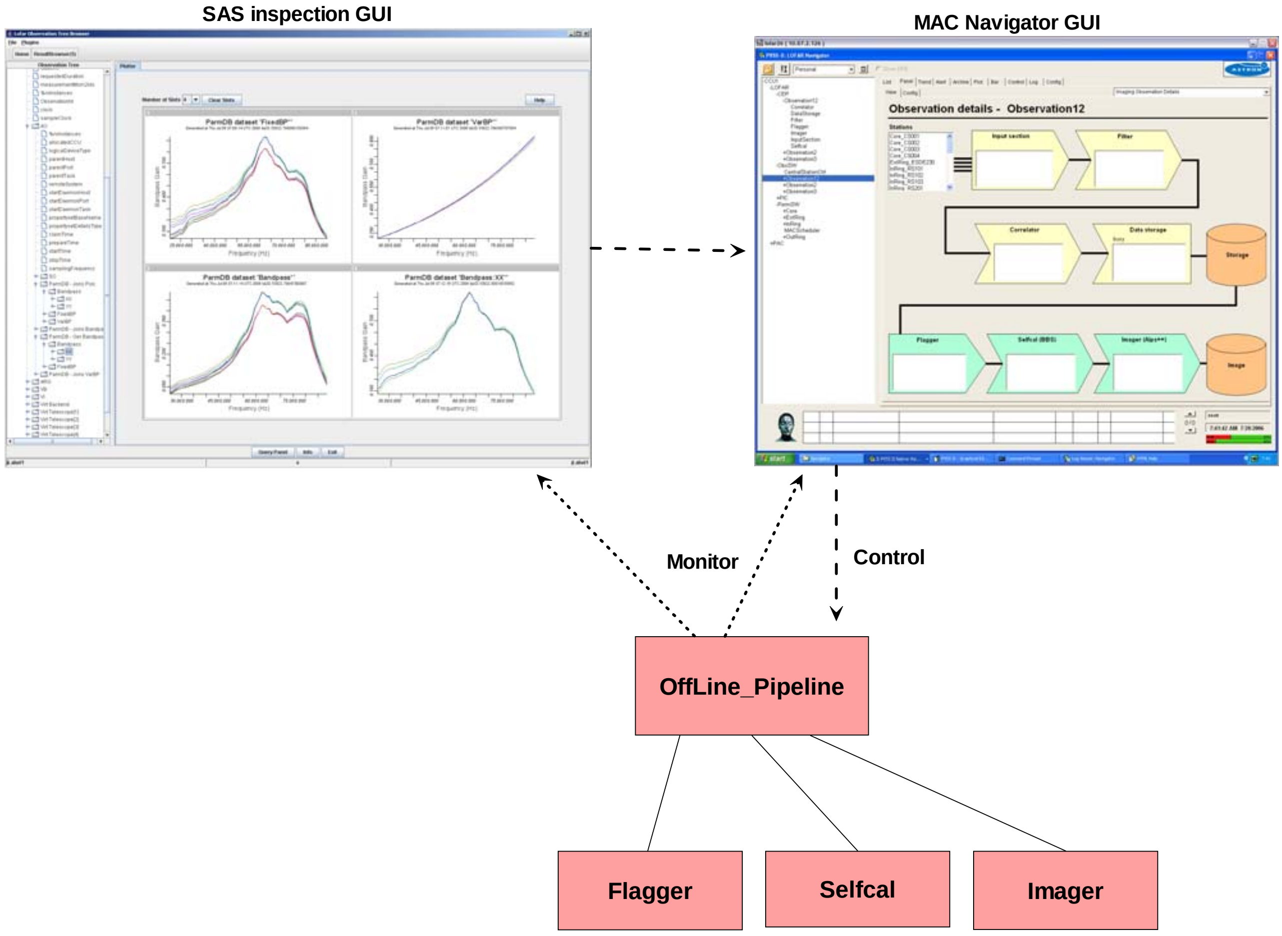

Figure 17 *Illustration of interactive control over a processing pipeline. In this example an imaging pipeline is inspected and controlled interactively by the SAS and MAC GUIs. As described in the text, this interactive usage is possible in the context of the full MAC control system.*

6.2.6 Data export and archiving

All data products generated in the on-line or off-line processing pipelines are stored in the CEP intermediate storage facility. From this location, data products are exported for further processing and storage either to a permanent LOFAR archive or to multiple archives located at the various Science Centres using the Surfnet6 infrastructure and GRID software. The Surfnet6 network is connected to the European Géant network and to the internet in general. These networks are used for user access to data products, either raw data, pre-processed data or final data products such as images.

7 Operational Modes and Performance

This Chapter describes the main LOFAR operational modes including descriptions of the typical usage and anticipated data products for each mode. Note that the flexibility of the LOFAR system is such that additional modes or variations on existing modes can be defined.

7.1 Observing Modes

The astronomical instrument LOFAR has four main modes of operation:

1) Aperture Synthesis Mode: In this mode, the data from the station beams are correlated to produce visibility data using the on-line systems. Depending on the specific needs of the observation, the data may be pre-calibrated and averaged in time and/or frequency to produce a more manageable data-rate for subsequent archiving and export. Calibration and further correction is typically handled by additional off-line processing. The most common use of this mode will be for imaging.

2) Tied Array Mode: For this mode, signals from antennas and/or stations are coherently or incoherently added to form sets of pencil beams, each generating a time-series voltage sum. Targeted pulsar observations and surveys will utilize this mode.

3) Transient Detection Mode: This mode monitors the stream of correlated visibility data in real-time for the presence of radio transients. Direct interrogation of the UV data or snapshot images may be queried to identify possible transient sources. Since it utilizes the correlated visibility data as a starting point, this mode can be run concurrently with aperture synthesis mode observations as well as in dedicated observations like the transient “all-sky” monitor.

4) Cosmic Ray Detection Mode: Like transient detection mode, this observation mode monitors the incoming data streams from stations in real time for the presence of radio pulses generated by cosmic ray air showers or even lightning. It differs from the previous transient mode in that it does not utilize correlated data, but rather the raw time series data from the stations. As with the previous real-time mode, this mode is also a good candidate for concurrent operation with other observing programs.

There are three key issues fundamental to any consideration of observing modes for LOFAR:

1) Sky Coverage: Remote station beam-forming results in the selection of one or more relatively small patches of the sky for further processing (generally less than a few percent of the field of view of individual receptors). Each core station will allow the construction of up to 24 digital beams, capable of filling a large part of the FOV.

2) Processing: Each observing mode may require its own unique processing chain (on-line and off-line). Users will have widely varying expectations from LOFAR with respect to data products, and widely varying capabilities for dealing with those data products. Available modes must cater to non-specialists who may only want a simple flux density for a source, for example. But also the highly experienced radio astronomers who wish to stretch LOFAR capabilities to their limits must be supported. Resource limitations will ultimately define the level of processing available to support a given mode.

3) Data Products: For many modes, the resulting data volumes and data flow rates will be very large. This fact leads to practical constraints on the type and amount of data a given observational mode

can produce as well as what final data products can reasonably be constructed, exported, or archived.

The range of difference in these characteristics naturally gives rise to a multiplicity of LOFAR observing modes and serves as a common point of comparison in the following discussion. In the remainder of this chapter, the current set of supported LOFAR observing modes are described. More detailed discussion of the processing pipelines and scientific applications are available in [4] and the project plans of the various Key Science Projects [14], [15], [16], [17].
.

7.2 Aperture Synthesis Modes

Generally, aperture synthesis mode observations are intended to result in an image of some portion of the sky, although for some applications the final set of data products may include the calibrated UV data as well. This mode is characterized by cross-correlation of signals from stations, with data analysis restricted to the field of view of the stations. Aperture synthesis imaging represents the coherent integration of signals at all points on the map, with the consequence that long wideband integrations can produce very low noise levels. For this reason, the aperture synthesis mode demands high precision calibration.

Different combinations of beams and bandwidth are possible based on the subband beamformers in the station digital processing system [3]. Those beams can be spatially independent, but are subbands within the same mode (see Section 3.2). For each remote station beam direction, a core station beam can be used as input for the station-core self-calibration process, which is used to calibrate for the ionosphere. Additional core station beams may be calculated and correlated with remote station beams for calibration on the remote station beam side-lobes. The station-core self-calibration application produces beamshape and phase corrections at 10 seconds time intervals. These parameters are used in the on-line pre-correlator correction function and are also available through the MAC database to other running pipeline components which require recent calibration information.

We discuss three major variants of the Aperture Synthesis mode, to highlight practical constraints imposed by ranges of observational parameters and goals. While most synthesis observations are expected to fall into these categories, a brief discussion of EoR imaging is presented, as the most demanding example of several anticipated special cases.

7.2.1 Standard Imaging

This mode is most analogous to synthesis imaging on current arrays. Typically, the goal will be to achieve high fidelity and low noise images of specific astronomical objects, using customized observing parameters. In this operating mode, station beams are transferred to the central processing facility where they are correlated to produce raw visibility data. This raw UV data is stored on the temporary storage cluster where further processing, primarily calibration, is handled off-line. Calibration will typically be an iterative process until the optimal level of correction is achieved. The final data products for this mode will typically include the calibrated UV data, possibly averaged in time and frequency, and the resulting image cubes for this data. Such averaging will reduce data volumes to manageable levels, and it will be possible to routinely export datasets to investigators for reduction and analysis at their Science Centre or suitable resources on the GRID.

For this type of observation, a full range of user interaction will be supported, from experienced users who require control over the calibration and imaging stages of data reduction, to mid-level users who do not wish to recalibrate, but may need to control imaging parameters, and finally to novice users who require only a fully processed image. The MAC system will provide complete control over all aspects of the calibration and imaging processing pipelines. For expert users, interactive control of this processing will be available using the SAS and MAC GUIs over the network.

This mode also will require medium to long-term storage of un-calibrated or partially calibrated data at the central processing facility, to allow reprocessing of data following detailed inspection of results by the user. The resulting storage and processing requirements may impose limits on the amount of such customized reprocessing which may be conducted in the early years of LOFAR operation.

7.2.2 Surveys

Imaging surveys represent a special case of the standard imaging mode. With its wide instantaneous field of view and multi-beaming features, the LOFAR system provides powerful survey capabilities. These capabilities combined with LOFAR's low-frequency sensitivity will make this mode both popular and potentially scientifically rich. Operationally, there is no difference between this mode and standard imaging in terms of the data transport, correlation, and subsequent offline calibration and imaging pipelines.
Survey mode does differ in terms of its expanded off-line processing to ensure uniform image quality, create all-sky image mosaics, and build complete catalogues of detected sources.

Because of the anticipated quantity of observations in this mode, a high degree of automation in the data reduction is essential, all of which will be conducted at the central processing facility. Data products will be in the form of images (either residual, or restored with GSM sources), and in the form of updates to the GSM. Storage of the prohibitively voluminous UV data will not be long-term.

7.2.3 Epoch of Reionisation (EOR)

Observations of the EOR represent an important special case of aperture synthesis mode and place some of the most stringent performance requirements in terms of storage and processing on the LOFAR system. Proof of a clear "reionisation" signal will come from the detection of significant changes in the spatial power spectrum of 'fluctuations', nominally in three dimensions, but principally in the frequency domain. This requires observations over a large frequency range (e.g. 120 – 200 MHz). For the EOR experiments, large parts of the sky have to be imaged for long periods.

The measurement consists of two parts, one to determine the properties of the discrete foreground source population, and one to measure the sum of the EOR signal and the foreground at high surface brightness sensitivity. The former uses a wide range of baseline lengths, while the latter uses only short baselines, most particularly those found within the core region. The most promising approach for LOFAR is to measure the spatial power spectrum of the EOR signal over as wide a range of spatial scales as possible. From these a large set of narrow-band images over a wide area of the sky and over a wide frequency range are produced. The observed spatial and spectral fluctuations would be analyzed via standard power spectrum analysis techniques. This approach would require an instantaneous bandwidth of at least 32MHz and full coverage of the primary beam of the High Band antenna, especially for baselines < 5km. Hence, at least 24 beams with 32 MHz bandwidth are necessary for the core stations.

Data will be voluminous due to the long observing durations, but must be archived for subsequent re-analysis as knowledge of the contaminating foreground improves. Care must therefore be taken during data averaging to manageable volumes in order to preserve sufficient information for such re-analysis. It is currently envisioned that after initial processing, the calibrated UV datasets will be averaged to resolutions of 10 sec and 10 kHz resulting in a factor of 100 reduction in size. These averaged datasets represent the basic starting point for all subsequent processing and analysis by the EoR project.

7.3 Tied Array modes

Operationally, tied array mode observations are distinguished from many of LOFAR's other operating modes in providing much more flexible beamforming capabilities. With the current system, it is possible to construct as many as 24 station beams for a total bandwidth of 32 MHz from each LOFAR core station. In tied array mode, various combinations of these station beams can be summed in phase to produce a potentially large set of small pencil beams or "tied array beams" (TABs) on the sky. The total number of tied array beams that can be formed will be dependent on the resources available in the central processing facility. Each TAB is capable of tracking sources making this mode highly efficient for scientific programs such as pulsar surveys.

This mode also differs from imaging modes in that the standard data product is not an image, but rather a dynamic spectrum associated with each TAB. The resulting voltage sum of each tied array beam may be presented to the user as multiple spectral subbands, or alternatively transformed back to the time domain by stitching the subbands together to produce a higher time resolution. The latter mode will be more commonly required, specifically for pulsar studies. The final data products will be raw time series data as well as time/frequency/polarization datacubes for each TAB.

It will also be possible to form tied array beams incoherently using combinations of the station beams. In this case, the incoming data for the station beams is squared before summing thereby removing information about the phase relationship between the beams. This processing produces a wider beam (i.e., the same size as the original station beam), but results in a decrease in gain compared to the coherently added tied array beam (the square root of the number of stations in the core). These incoherently summed beams provide a highly efficient way to conduct wide-area pulsar surveys.

Generally, tied array observations are compatible with simultaneous aperture synthesis mode and transient mode observations, with specific restrictions noted below. We discuss two important sub-modes for tied array observations below.

7.3.1 Known pulsar mode

For observations of known pulsars, as well as some classes of transients such as planets, a tied array mode observation can be used to obtain dynamic spectra at high time resolution. In this case the station data is transferred to the CEP where it is beamformed to acquire and track the identified source. The resulting raw time series data can be stored at this stage for more detailed off-line analysis or processed on-line.

For all pulsar observations, a de-dispersion correction will be part of the standard processing pipeline. These corrections and subsequent processing steps will be implemented as an on-line pipeline to take advantage of the CEP computational capabilities. An off-line version of this pipeline will also be available for processing pulsar observations in the event that CEP resources are not available at the time of observation.

This submode forms tied array beams using only signals from within the core region of 2 km diameter. In this region, 24 beams of 32 MHz instantaneous bandwidth are produced (more beams can be generated when bandwidth is exchanged). This provides a high degree of flexibility. Still a conflict with the aperture synthesis observing mode is present when observing simultaneously. High time resolution monitoring of a large numbers of objects (e.g. pulsars, planets, flare stars, or exotic binaries) are possible in this mode.

7.3.2 Pulsar survey mode

Unlike systematic studies of known pulsars, which will primarily be conducted using the previous mode, searches for new pulsars will be conducted in one of two ways. First, for coherent pulsar searches, tied array beams will be constructed in the standard manner. The efficiency of this search program will be determined by the number of available TABs which will depend on system resources. As opposed to known pulsars for which the appropriate dispersion corrections will be known in advance, pulsar search processing will need to correct the incoming time series data for a range of “trial” Dispersion Measures (DM). The resulting grid of corrected data must then be searched for candidate pulsars. This dispersion measure search is the major processing load for this mode. It is currently planned to implement this DM search as part of the on-line processing on CEP. Offl-ine processing may also be possible but would place additional requirements on the storage capacity of the temporary storage cluster.

Incoherent pulsar surveys differ only in the manner in which the TABs are formed. The post-beam-forming processing including the dispersion measure search will be the same. We note that the same dispersion measure search will be required to achieve monitoring of other transient sources at the shortest timescales.

7.4 Transient Detection modes

Unlike those discussed previously, the transient detection mode consists of a real-time pipeline to monitor the incoming stream of correlated data for both known variable sources as well as new, unknown radio transients. When a new transient source is detected or a known monitoring source undergoes a rapid state change, this mode is designed to respond on short timescales (down to ~1 sec) and generate triggers to the LOFAR system to initiate a number of possible responses. These responses might include switching to a different, targeted observational mode, adjusting the subband selection for more optimal frequency coverage, initiating a dump of the transient buffer data for more detailed processing, or sending an alert to other observatories to initiate coordinated observations. Similarly, LOFAR will also be capable of responding to external triggers from other observatories to initiate coordinated monitoring of transient sources.

The real-time nature of this observing mode requires that the associated processing pipeline run as part of the on-line system on the CEP hardware. In principle, this mode can be operated concurrently with other observations in a “piggy-back” mode by querying the streaming visibilities from those observations. We discuss two dedicated transient modes below although in principle the associated processing chains are quite similar. In general, the real-time nature of this mode will preclude human intervention; therefore, the processing must be fully automated. The data products associated with this mode will consist of snapshot images integrated over different time-scales as well as UV datasets suitably averaged in frequency. Full time resolution will be preserved in order to allow more careful off-line processing to study the temporal behaviour of any detected transients.

The primary performance criterion for this observation mode is the response time to detected transient sources, whether newly discovered or a known source entering a new state. The time required to transport, process, and detect a transient event introduces latency in the LOFAR system’s ability to respond to such an event. Consequently, for trigger follow-ups, data buffering is needed. The most flexibility is offered by buffering before the beamformers. If fast re-programming of the beamformers is available, an ideal follow-up observation within the observed frequency range can be performed. At the time of writing, a number of different methods for reducing the latency of the LOFAR system are being considered.

7.4.1 Known Transient Monitoring Mode

Although primarily intended as a means of detecting new transients, the same mode can be used in targeted observations to monitor known variable sources. For example, this mode might be used to monitor the radio emission from a known flaring X-ray binary star with high time and/or spatial resolution perhaps as part of a coordinated multi-wavelength campaign. The incoming correlated data stream is monitored for variations and, based on pre-defined heuristics for identifying types of source variability the MAC system can generate triggers to initiate appropriate actions.

In order to keep up with the data flow, the streaming visibility data will not undergo the sort of rigorous, iterative self-calibration found in the standard LOFAR imaging modes. Rather a simple correction based on the most recent calibration available from the MAC database will be applied. As mentioned above, these calibration solutions will be available at roughly 10 sec intervals.

The large RAM buffer of the transient buffer boards (TBBs) allows for the possibility of a “playback” mode in the monitoring chain. Following the generation of a trigger event, due to a sudden increase in source intensity for example, the contents of these buffers can be frozen and routed back to the central processing facility. At CEP, the data can either be stored for more detailed off-line processing or they can be routed through the currently running on-line processing. These buffered data retain the full field of view of the antennas, and provide a capability to perform post-facto observing in any direction, should the need arise.

7.4.2 Radio Sky Monitor Mode

In the Radio Sky Monitor (RSM) mode, multiple beams from the LOFAR core stations will be used to cover a large fraction of the field of view. At low frequency, this coverage will allow a significant fraction of the entire sky to be monitored for transient low frequency radio sources. The range of timescales which can be examined for variability are determined by the length of a given observation at the upper end and the current correlator dump time of ~1 second at the lower bound. The total RSM field-of-view and minimum variability time-scale probed will be determined by the available bandwidth for data transport and the available computing power for processing.

Direct interrogation of the incoming UV data stream or snapshot images will be queried to identify possible transient sources. To detect transients on different timescales, (and with different flux densities), detection runs will performed on logarithmically-spaced intervals in integration time (1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000 seconds). In order to benefit from the increased UV-coverage as integration time increases, these images will be constructed from the appropriate calibrated UV data rather than simply stacking a series of 1 sec images. A copy of the compressed UV data will also be archived for later reference or possible re-analysis.

In addition, to detect short timescale transients the effects of dispersion in the interstellar medium must be corrected. This process will not affect long timescale radio transients since the timescales are sufficiently long that there is no “burst” as such to disperse. Consequently for shorter timescales, maps and residual images will need to be computed at a range of dispersion measure corrections before detection.

7.5 Cosmic Ray Detection Modes

The electronic agility of LOFAR combined with the wide-field nature of its receivers makes the instrument uniquely powerful for the detection and study of the radio emission from extensive air showers initiated by Cosmic Rays. The potential of LOFAR in this area is exploited using multiple subsystems, namely the transient buffer and the air shower pulse detection system. Radio pulses generated by cosmic ray air-showers, with a time scale of order 10 ns can be detected using a system based on coincidental total power pulses localized over an air-shower footprint of a few hundred meters. Upon a trigger from this detection system, the relevant portion of the transient buffer is frozen and downloaded to the central processing facility for further processing or shipment to investigators. The geometrical distribution of the dipoles can be exploited to obtain time resolutions lower than the sample time of the digitizer.

A number of different cosmic ray detection sub-modes are supported. They differ primarily in the manner in which detections are identified and triggers are generated. These behaviors are in turn tied to the energy of the cosmic ray they are designed to detect. For example, since the strength of the associated radio pulse depends on the energy of the primary cosmic ray, for lower energy events the signal to noise ratio of the pulse will be too low to detect in the data stream from an individual antenna. In these cases, a tied array beam is formed before pulse detection. Once a pulse is detected and a trigger is generated, all these modes result in a dump of the data in the transient buffer. This data is stored for subsequent off-line processing.

All cosmic ray detection modes place a strong constraint on the response time of the LOFAR system. The system must be able to process a detected pulse and dump the contents of the transient buffer within a time interval which is smaller than the length of the buffer itself, otherwise the data from the event will be lost. The operational details for the various sub-modes are discussed individually.

7.5.1 HECR mode

For High Energy Cosmic Rays (HECR), the radio signal produced by an extensive air shower is too weak to be detected in the raw, time-series data of an individual dipole. For this reason, the signal detection is performed on a time-series obtained from the streaming station beam data. This detection mode can consequently operate very naturally concurrent with other observations modes. In order to produce this time-series from the beamformed station data, the effect of the polyphase filter bank must be inverted. The resulting time-series per beam direction can then be scanned for pulses. If such a pulse is detected, a freeze of the TBB data is initiated.

In Figure 18 the scenario for a HECR detection is depicted. CEP is for this application configured (partly) in the HECR mode and receives the beam data from the stations. If a pulse is detected in the time-series than a trigger (1) will be send to MAC. MAC will take the initiative to freeze buffers in one or more stations by the freeze (2) command. When the buffers are frozen, CEP is set ready by MAC to receive the buffered data (3). After that, the stations are controlled to send the required buffered data to CEP (4). Finally, the data is sent from the station buffers to CEP (5).

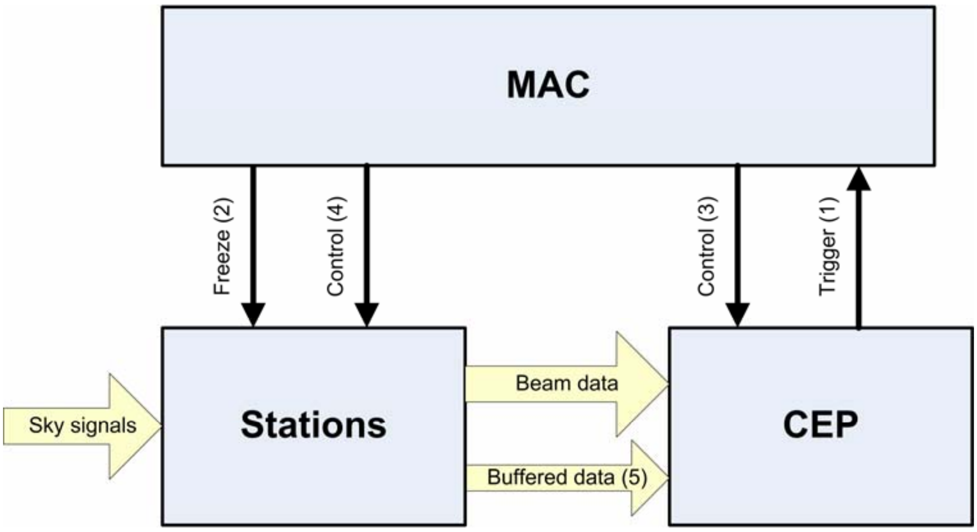

Figure 18 *HECR trigger scenario at system level (only the relevant signals are shown)*

7.5.2 UHEP mode

The Ultra High Energy Particle (UHEP) mode is designed to search for radio peaks originating from the interaction of the highest energy particles with the lunar regolith. In order to detect this radio emission, parts of the LOFAR array are phased up to form beams towards the surface of the moon. The result tied array beams are then passed through a de-dispersion step to correct for the dispersion due to passage through the Earth's atmosphere. After de-dispersion, the effects of the polyphase filter bank are again inverted and the resulting time-series per beam direction is scanned for pulses. If a pulse is detected, a freeze of the TBB data is once again initiated. The HECR trigger scenario depicted in Figure 18 applies also for the UHEP mode.

7.5.3 VHECR mode

In Very High-Energy Cosmic Ray (VHECR) detection mode, the radio signal produced by an extensive air shower is strong enough to be detected in the raw time-series of individual dipoles. Consequently, the first two stages of the detection process are carried out at the station level. First, a pulse detection algorithm is run on the TBB themselves. Signals from those station dipoles reporting detections are then combined at the station LCU to decide whether or not a pulse has been detected. The results of these station level detections are then transported to the MAC system where MAC performs additional coincidence checks and initiates dumps of the relevant transient buffers.

The necessary data transport rates and resulting data volumes associated with this detection mode could in principle be very large since a sufficiently large air shower could potentially trigger detections in a large fraction of the antennas over the whole LOFAR array.

In Figure 19 the VHECR trigger scenario is shown. For this mode the pulse is detected at the station level and that generates a trigger (1) to MAC. MAC initiates a freeze (2) on the storage boards, dependent on how many antennas a trigger was generated. Also the freeze can be local to one station or involve multiple stations. CEP is set ready by the control signal (3) of MAC to receive the buffered data. MAC controls the station (4) and identifies which data should be sent (5) to CEP.

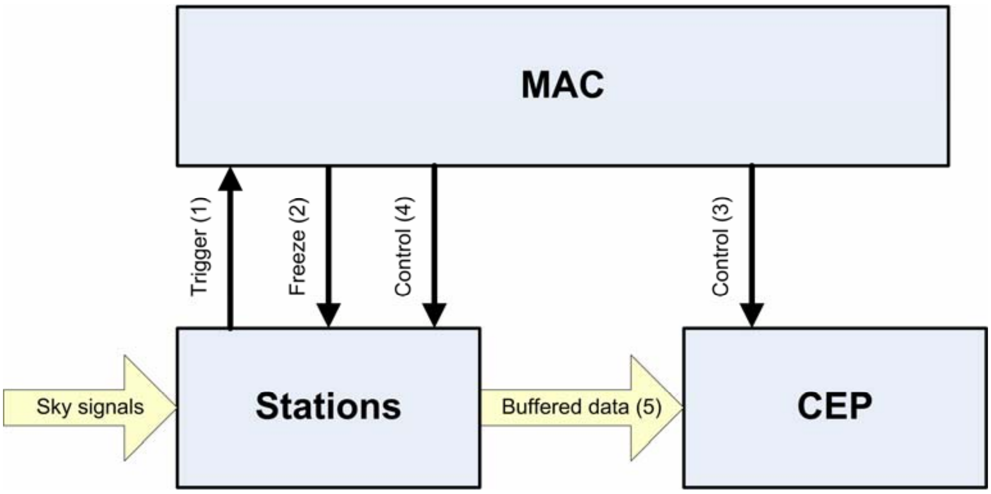

Figure 19 *VHECR trigger scenario at system level (only the relevant signals are shown)*

8 Open ends

8.1 Core stations and processing

For the EOR and transient detection modes, 24 beams in the core stations are required as mentioned in Section 2.2. Normally the remote stations deliver only 1 beam of 32 MHz in bandwidth. Increasing the number of beams of the core stations has a significant impact on the system architecture. The current development effort was focused on the remote stations. These remote stations can be re-used for the core as well with taking into account an upgrade. The way in which a remote station can be upgraded to a core station is described in [3]. Note that there is not yet made an architectural choice where to pre-process the 24 beams. The reason for this is that first, experience needs to be gained and the boundaries of the current system need to be explored.

It should be noted that the 24 beams mode will only be used for the core stations and cannot be used simultaneously with an observation running on the remote stations.

8.1.1 The issue

Normally the system is able to transport at most a single 32 MHz beam from 77 stations to the central systems. The 24 beams mode will re-distribute the total I/O and CEP processing capacity so that it is fully used by the 32 core stations. In order to accommodate the 24 beams, a factor 4 reduction in data-rate will be achieved by bit-reduction: the 16 bit complex words will be replaced by 4 bit versions. This should be preferably done by selecting only the RFI free channels, however, since the second filterbank (which provides the ~1 kHz spectral resolution) is currently implemented in the CEP subsystem in Groningen, further modifications will be required.

Two options are available:

- implement the channel filter bank in the core stations (note that the number of channel filter banks for the 24 beams mode is 24 times the required number of channel filter banks in the standard mode)
- implement the bit reduction at the stations after the subband filter bank

Both are discussed in the next sections and further referred to as respectively the performance impact minimized option and architectural impact minimized option.

8.1.2 Performance impact minimized

In the performance neutral option the channel filterbank should be implemented in the core stations. As a consequence the delay tracking (Figure 8) should be implemented as well in the core stations. This adds significantly to the complexity from an architectural point of view.

After the channel filter bank RFI detection is necessary to indicate which channels are not worthwhile to process any further because the required dynamic range is exceeding 4 bits. The channels which will be transported to the central systems should be accordingly compressed in the number of bits. The design of this option is presented in [29].

8.1.3 Architectural impact minimized

An architectural neutral option is to keep the channel filter in the central systems and accomplish a data reduction on subbands instead of channels. The consequence is that narrow band RFI signals within a subband make the complete subband useless to process further. Hence, a 24 beam observation can never be done close to RFI signals. Based on the data which is currently generated in the field the consequence of the last statement will be investigated.

The central systems are able to cope with 24 more filterbanks per core station [28].

9 References

[1] M.P. van Haarlem, H. Kollen (editor), LOFAR System Requirement Specification, LOFAR-ASTRON-SRS-001, March 2007
[2] R. Braun, A. Huijgen, LOFAR Configuration Comparisons, February 2006
[3] A.W. Gunst, G.W. Schoonderbeek, LOFAR Station Architectural Design Document, LOFAR-ASTRON-ADD-013, March 2007
[4] A. Latour, K. v.d. Schaaf, LOFAR Central Processing Facility Architecture description, LOFAR-ASTRON-ADD-012, March 2007
[5] G.W. Kant, A.W. Gunst, HBA ADD, LOFAR-ASTRON-ADD-016, March 2007
[6] S. Cadot, LOFAR System Health Management ADD, LOFAR-SHM-ADD-001, October 2005
[7] D.H.P. Maat, R.B. gloudemans, LOFAR-WAN ADD, LOFAR-ASTRON-ADD-017, March 2007
[8] R. Overeem, Specification Administration & Scheduling Architectural Design Document, LOFAR-ASTRON-SDD-041, March 2007
[9] R. Overeem, Monitor And Control Architectural Design Document, LOFAR-ASTRON-ADD-005, March 2007
[10] S.J. Wijnholds, Study on LOFAR Core Calibratability Based on Source Statistics, LOFAR-ASTRON-MEM-205, April 2006.
[11] S.J. Wijnholds, Pre-correlation Signal Processing Baseine, LOFAR-ASTRON-MEM-105, August 2003
[12] M.A. Brentjes, A. Gerdes, Decorrelation effects in delay compensation, LOFAR-ASTRON-MEM-158, August 2004
[13] B. Razavi. RF Microelectronics. Prentice Hall, Upper Saddle River, NJ, 1998.
[14] G. de Bruyn, S. Zaroubi, L. Koopmans, LOFAR Key Project: The Epoch of Reionization, May 2006
[15] R. Braun, H. Falcke, R. Fender, J.-M. Griessmeier, P. Jonker, M. Kramer, S. Markoff, J. Miller-Jones, R. Osten, H. Spreeuw, B. Stappers, R. Wijnands, R. Wijers, M. Wise, P. Zarka, The LOFAR Transients Key Project, May 2006
[16] H.J.A. Rottgering, P.D. Barthel, M. van Haarlem, G.K. Miley, N. Mohan, R. Morganti, I. Snellen, LOFAR Surveys of the Radio Sky, May 2006
[17] J. Bacelar, L. Bähren, S. Buitink, H. Falcke, A. Haungs, A. Horneffer, T. Huege, J. Kuijpers, S. Lafebre, S. Nehls, A. Nigl, J. Petrović, O. Scholten, K. Singh, C. Timmermans, A.M. van den Berg, Project Plan LOFAR/CR, May 2006
[18] W.A. Baan, A.J. Boonstra, Site Characteristics for LOFAR, Drenthe Site, The Netherlands, LOFAR-ASTRON-MEM-023, October 2001
[19] A. J. Boonstra, LOFAR RFI site studies for Drenthe, LOFAR-ASTRON-RPT-097,March 2003
[20] W.A. van Cappellen, LBA Station Beam Simulations, LOFAR-ASTRON-RPT-099, February 2006
[21] W.A. van Cappellen, Addition to LBA Station Beam Simulations, LOFAR-ASTRON-RPT-100, March 2006
[22] W.A. van Cappellen, Addition to LBA Station Beam Simulations Irregular configurations, LOFAR-ASTRON-RPT-101, March 2006
[23] W.A. van Cappellen, Correction of LBA Station Beam Simulations, LOFAR-ASTRON-RPT-102, March 2006
[24] W.A. van Cappellen, Update of LBA Station Beam Simulations, LOFAR-ASTRON-RPT-103, May 2006
[25] S.J. Wijnholds, Station Calibration Pitfalls & Possibilities 2006, LOFAR-ASTRON-MEM-217, October 2006
[26] K. v.d. Schaaf, R.J. Nijboer, LOFAR Calibration Implementation, LOFAR-ASTRON-SDD-050, version 2, March 2007
[27] G.M. Loose, J.E. van Zwieten, Blackboard Self-calibration (BBS), Software Design Document, LOFAR-ASTRON-SDD-053, March 2007
[28] J. Romein, On-Line Application Processing (OLAP), LOFAR-ASTRON-SDD-036, March 2006
[29] A.W. Gunst, Core Station Baseline, LOFAR-ASTRON-MEM, July 2006

[30] A.J. Boonstra, M. Tanigawa, LOFAR Interference Mitigation Approach and Measurement Results, LOFAR-ASTRON-RPT-104-1.0, October 2006
[31] A.J. Boonstra, S. van der Tol, "Spatial Filtering of Interfering Signals at the Initial Low Frequency Array (LOFAR) Phased Array Test Station", Radio Science, vol. 40, 2005. (downloadable from http://ens.ewi.tudelft.nl/Research/array/VICI/publications.php)
[32] A.J. Boonstra and S.J. Wijnholds, "Implementation of the LOFAR Interference Mitigation strategy, considerations and tradeoffs", URSI, New Delhi, India, 23-29 October 2005, LOFAR-ASTRON-PUB-012
[33] W. Lubberhuizen, "RSP Firmware Functional Specification", LOFAR-ASTRON-SDD-001, 15 September 2005.
[34] A.J. Boonstra, "Radio Frequency Interference Mitigation in Radio Astronomy", Thesis, Delft University of Technology, June 14, 2005, downloadable from http://www.darenet.nl/
[35] J. Raza, A. Boonstra and A. van der Veen, "Spatial filtering of RF interference in radio astronomy", IEEE Signal Processing Letters, vol. 9, no. 3, Mar. 2002.
[36] S.W. Ellingson, J. Bunton, and J.F. Bell, "Removal of the GLONASS C/A Signal from OH Spectral Line Observations Using a Parametric Modelling Technique", *Astrophysical J. Supp. Ser.*, 135(1):87-93 (July 2001).
[37] A.J. Boonstra, "LOFAR RFI site studies for Drenthe", 6 March 2003, LOFAR-ASTRON-RPT-097.
[38] A.J.Boonstra and R.Millenaar, "LOFAR monitoring, Schoonloo site survey, site visit: May 25/26, 2004", LOFAR-ASTRON-RPT-87-1.0, July 27, 2006.
[39] B.Woestenburg, H. van der Marel, and A.G. de Bruyn, "LFFE Final Acceptance Report", 16 March 2006, ASTRON- RP-102.
[40] J.E. Noordam, LOFAR Calibration Framework, LOFAR-ASTRON-ADD-015, 2006
[41] A.G. de Bruyn, R.J. Nijboer, LOFAR Calibration and Commissioning Plan, LOFAR-ASTRON-PLN-033, March 2007
[42] J.W. Romein et al, Astronomical Real-Time Streaming Signal Processing on a Blue Gene/L Supercomputer, proceeding of 18th ACM Symposium on Parallelism in Algorithms and Architectures, 2006
[43] K. v.d. Schaaf, LOFAR Software architecture, LOFAR-ASTRON-SDD-049, 2006
[44] A. v. Houwelingen, G.W. Schoonderbeek, Detailed Description of Clock and Sync, LOFAR-RPT-057, March 2007